

**AD A 055632**

**FOR FURTHER TRAINING**

**P-77-5**

*[Handwritten signature]*

*[Handwritten signature]*

## **HANDBOOK FOR THE DEVELOPMENT OF SKILL QUALIFICATION TESTS**

**by**

**William C. Osborn, Roy C. Campbell and J. Patrick Ford**  
**HUMAN RESOURCES RESEARCH ORGANIZATION**

**and**

**Stephen F. Hirshfeld and Milton H. Maier**  
**ARMY RESEARCH INSTITUTE FOR THE**  
**BEHAVIORAL AND SOCIAL SCIENCES**

**NOVEMBER 1977**

**Contract DAHC-19-76-C-0044**

**Prepared for**



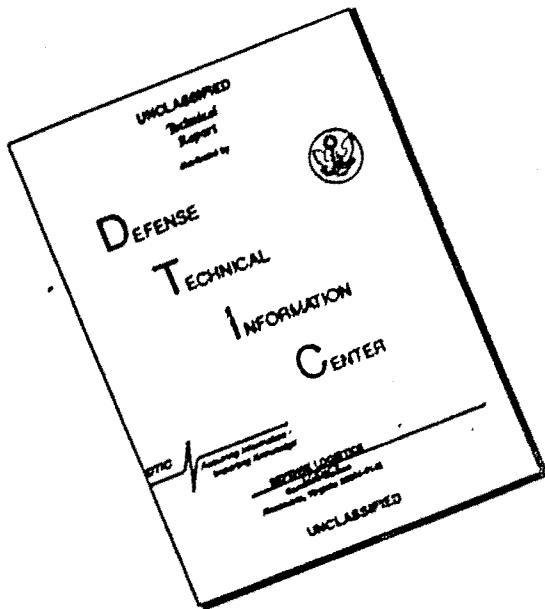
**U.S. ARMY RESEARCH INSTITUTE  
for the BEHAVIORAL and SOCIAL SCIENCES  
5001 Eisenhower Avenue  
Alexandria, Virginia 22333**



**Approved for public release; distribution unlimited.**

**78 06 22 013**

# **DISCLAIMER NOTICE**



**THIS DOCUMENT IS BEST  
QUALITY AVAILABLE. THE COPY  
FURNISHED TO DTIC CONTAINED  
A SIGNIFICANT NUMBER OF  
PAGES WHICH DO NOT  
REPRODUCE LEGIBLY.**

**U. S. ARMY RESEARCH INSTITUTE  
FOR THE BEHAVIORAL AND SOCIAL SCIENCES**  
**A Field Operating Agency under the Jurisdiction of the  
Deputy Chief of Staff for Personnel**

---

**J. E. UHLANER**  
**Technical Director**

**W. C. MAUS**  
**COL, GS**  
**Commander**

---

Research accomplished  
under contract to the Department of the Army

Human Resources Research Organization

**NOTICES**

**DISTRIBUTION:** Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U. S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-P, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.

**FINAL DISPOSITION:** This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Research Institute for the Behavioral and Social Sciences.

**NOTE:** The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

14 HUMRRO-FR-WD-KY-77-1

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

18 ARI REPORT DOCUMENTATION PAGE

REPORT NUMBER

19 P-77-5 ✓

2. GOVT ACCESSION NO.

READ INSTRUCTIONS  
BEFORE COMPLETING FORM

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and subtitle)

6 HANDBOOK FOR DEVELOPMENT OF SKILL QUALIFICATION TESTS

AUTHOR

10 William Osborn, Roy C. Campbell, J. Patrick Ford,  
Stephen F. Hirshfeld Milton H. Maier

5. PERFORMING ORGANIZATION NAME AND ADDRESS

Human Resources Research Organization ✓  
300 North Washington Street  
Alexandria, Virginia 22314

11. CONTROLLING OFFICE NAME AND ADDRESS

U.S. Army Training and Doctrine Command (TRADOC)  
Individual Training Evaluation Directorate  
Fort Eustis, Virginia 23604

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

U.S. Army Research Institute for the Behavioral  
and Social Sciences PERI-II  
5001 Eisenhower Ave., Alexandria, VA 22333

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Research and development performed by HumRRO Western Division, Fort Knox  
Office, Fort Knox, Kentucky, monitored technically by personnel of the  
Individual Training and Skill Evaluation Technical Area of ARI.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Performance Testing	Skill Qualification Test (SQT)
Job Proficiency Testing	Criterion Reference Testing
Personnel Evaluation	Test Development

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Skill Qualification Tests (SQT) are the quality control instruments for  
the Army's new Enlisted Personnel Management System. The tests, used to  
evaluate soldier's job mastery, are developed by personnel in Army agencies  
responsible for the various job specialties. Since the SQT entails a new  
performance-oriented approach to proficiency testing, agency personnel must  
be trained in the techniques of SQT development. The present handbook was  
prepared for use by SQT developers in the Army. The handbook covers both  
technical and administrative procedures to follow in preparing a field-

DD FORM 1 JAN 73 1473 / EDITION OF 1 NOV 68 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

78 06 22 013  
405260

(over)

*cont*

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. Attested SQT. Included are techniques and procedures pertaining to: which job tasks to test; what method of testing to use; how to develop a hands-on test; how to develop a performance-oriented written test; how to validate tests; and how to prepare guidance for test administration.

*[Signature]*

The Army Training and Doctrine Command (TRADOC) is distributing the handbook for use by Test Development Activities.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

## TABLE OF CONTENTS

	Page
<b>Chapter 1. INTRODUCTION</b>	<b>1-1</b>
<b>Section I. General</b>	<b>1-1</b>
<b>Section II. How to use this Handbook</b>	<b>1-3</b>
<b>Section III. Definitions</b>	<b>1-4</b>
<b>Section IV. SQT Development Process Overview</b>	<b>1-10</b>
<b>Chapter 2. SELECT TASKS FOR TESTING</b>	<b>2-1</b>
<b>Chapter 3. REVIEW TASK ANALYSIS</b>	<b>3-1</b>
<b>Chapter 4. ALLOCATE TASKS TO SQT COMPONENTS</b>	<b>4-1</b>
<b>Section I. Task Allocation</b>	<b>4-1</b>
<b>Section II. Prepare SQT Plan</b>	<b>4-5</b>
<b>Chapter 5. CONSTRUCT HANDS-ON COMPONENT SCORABLE UNIT</b>	<b>5-1</b>
<b>Section I. Determine Product or Process</b>	<b>5-2</b>
<b>Section II. Specify Performance Measures: Process</b>	<b>5-6</b>
<b>Section III. Specify Performance Measures: Product</b>	<b>5-10</b>
<b>Section IV. Determine Test Conditions</b>	<b>5-15</b>
<b>Section V. Prepare Examinee Instructions</b>	<b>5-18</b>
<b>Section VI. Prepare Scorer Instructions</b>	<b>5-22</b>
<b>Chapter 6. VALIDATE HANDS-ON COMPONENT</b>	<b>6-1</b>
<b>Section I. Try out Hands-On Component SU</b>	<b>6-1</b>
<b>Section II. Validate Hands-On SU</b>	<b>6-7</b>
<b>Section III. Conduct Field Tryout of HOC</b>	<b>6-11</b>



	Page
<b>Chapter 7. CONSTRUCT WRITTEN SCORABLE UNIT</b>	<b>7-1</b>
Section I. Determine Written Performance or Performance-Based	7-3
Section II. Develop Written Performance Item	7-7
Section III. Develop Performance-Based Written Scorable Unit	7-13
<b>Chapter 8. VALIDATE WRITTEN COMPONENT</b>	<b>8-1</b>
Section I. Try Out Written SU	8-1
Section II. Try Out Written Component	8-11
Section III. Validate Two Written SU Against Hands-On	8-13
Section IV. Select Validation Option for WC	8-16
Section V. Validate Written SU - Self Ratings	8-19
Section VI. Validate Written SU - Supervisor Rating	8-31
Section VII. Validate Written SU - Panel-of-Experts	8-39
<b>Chapter 9. DEVELOP PERFORMANCE CERTIFICATION (PCC) SU</b>	<b>9-1</b>
<b>Chapter 10. PREPARE SQT NOTICE</b>	<b>10-1</b>
<b>APPENDIX A. PERFORMANCE CERTIFICATION COMPONENT SQT 2, MOSC 16J10 SCORING INSTRUCTIONS TO SUPERVISORS</b>	<b>A-1</b>
B. SQT NOTICE	B-1

	Page
<b>FIGURES</b>	
<b>Figure 1-1. Overview of SQT Development Process</b>	<b>1-9</b>
<b>2-1. Select Tasks for Testing</b>	<b>2-3</b>
<b>3-1. Review Task Analysis</b>	<b>3-3</b>
<b>3-2. Sample Questionnaire</b>	<b>3-9</b>
<b>4-1. Allocate Tasks to SQT Components</b>	<b>4-0</b>
<b>4-2. SQT Plan Outline</b>	<b>4-6</b>
<b>5-1. Hands-On Component Overview</b>	<b>5-0</b>
<b>5-2. Determine Product or Process</b>	<b>5-3</b>
<b>5-3. Specify Performance Measures: Process</b>	<b>5-7</b>
<b>5-4. Specify Performance Measures: Product</b>	<b>5-11</b>
<b>5-5. Determine Test Conditions</b>	<b>5-14</b>
<b>5-6. Prepare Examinee Instructions</b>	<b>5-19</b>
<b>5-7. Prepare Scorer Instructions</b>	<b>5-23</b>
<b>6-1. Try out Hands-On Component SU</b>	<b>6-0</b>
<b>6-2. Question Sheet</b>	<b>6-5</b>
<b>6-3. Validate Hands-On SU</b>	<b>6-6</b>
<b>6-4. Conduct Field Tryout of HOC</b>	<b>6-12</b>
<b>6-5. Examinee Questions for Hands-On Scorable Units</b>	<b>6-14</b>
<b>7-1. Develop Written Scorable Unit Overview</b>	<b>7-2</b>
<b>7-2. Determine Written Performance or Performance-Based</b>	<b>7-4</b>
<b>7-3. Develop Written Performance SU</b>	<b>7-8</b>
<b>7-4. Develop Performance-Based Items Items</b>	<b>7-14</b>
<b>7-5. Examples of Written-Performance Items</b>	<b>7-23</b>

	Page
<b>FIGURES</b>	
Figure 8-1. Try out Written Component SU	8-2
8-2. Subject Matter Expert Questionnaire	8-5
8-3. Try out Written Component	8-10
8-4. Validate Two Written SU Against Hands-On	8-12
8-5. Select Validation Option for WC	8-17
8-6. Validate Written SU - Self Ratings	8-20
8-7. Question Sheet 2	8-22
8-8. Item Analysis Form	8-24
8-9. Scorable Unit Analysis Sheet - Four or Less Items	8-28
8-10. Scorable Unit Form	8-29
8-11. Validate Written Component-Supervisor Rating	8-32
8-12. Validate Written SU - Panel of Experts	8-38
9-1. Develop Performance Certification (PCC) SU	9-0
10-1. Prepare SQT Notice	10-0

HANDBOOK FOR THE DEVELOPMENT OF SKILL QUALIFICATION TESTS

CHAPTER 1

INTRODUCTION

Section I. GENERAL

1-1. PURPOSE. The purpose of this handbook is to provide general guidance on construction and validation of the Skill Qualification Test (SQT) to personnel responsible for preparing SQT's for the Active Army, National Guard, and Army Reserve.

1-2. INTRODUCTION.

a. The Enlisted Personnel Management System (EPMS) is designed to meet requirements of the US Army to integrate various management and training subsystems of the Army. EPMS features training and education as a progressive career process dependent upon a comprehensive proficiency training and testing program to evaluate competence of enlisted soldiers at all skill levels.

The proficiency test program developed to verify competence is the SQT. The goal of each SQT is to provide an equitable, reliable, and relevant means of determining job proficiency of enlisted soldiers.

b. Five skill levels under EMPS standardize the relationship between a soldier's grade and his skill level. They are -

<u>Grade</u>	<u>Skill Level (SL)</u>
E-8, E-9	SL5
E-7	SL4
E-6	SL3
E-5	SL2
E-1 thru E-4	SL1

1-3. TOTAL PLAN. The SQT is part of the total EPMS plan which includes development of an Outline Individual Training Plan (OITP), construction of Soldier's Manuals and initiation of resident and extension training.

1-4. SEQUENCE. After reorganization of career management fields by USA MILPERCEN, proponent service schools must analyze each job or Military Occupational Specialty (MOS) to determine the tasks critical to duty performance at each skill level. That constrained task list becomes the basis for all subsequent EMPS actions.

The OITP evolves from the list of critical tasks as each proponent school decides on the training site for each task of a given MOS. The OITP delineates which will be mastered initially by extension training, and which will be reserved for initial mastery in the service schools or training centers.

Since these decisions bear directly on the kinds of resources required by schools to provide training support to the MOS, the OITP will be forwarded to HQ, U.S. Army Training and Doctrine Command, (TRADOC) for approval not later than nine months prior to the Career Management Field (CMF) effective date.

The CMF effective date is defined as the date upon which training will begin for that CMF under the EPMS system. Guidance for construction of the Individual Training Plan is provided in TRADOC Regulation 351-3.

1-5. SOLDIER'S MANUAL. Construction of the Soldier's Manual follows formulation of the OITP. Draft Soldier's Manuals will be forwarded to TRADOC for review and approval 18 months before the first day of the test quarter for the first SQT. Subsequent to TRADOC approval, camera-ready copy of the Soldier's Manual will be forwarded to TRADOC 10 months prior to the first day of the test quarter for the first SQT. This will facilitate printing and distribution of Soldier's Manuals to MOS incumbents at least six months prior to the SQT.

1-6. TESTS. Skill Qualification Tests will be administered to everyone in a CMF at any time during a test quarter. MOS incumbents will be evaluated only by EER until an SQT is available. The SQT plan for each MOS will be forwarded to HQ, TRADOC, 18 months before the first SQT test quarter. The final draft SQT and the SQT Notice will be forwarded nine months before the first SQT test quarter.

1-7. TASKING LETTERS. Tasking letters established individual milestone schedules have been worked out with each proponent TDA. Check the tasking letter for your TDA to find the overall development plan for your SQT.

## Section II. HOW TO USE THIS HANDBOOK

1 -8. STEPS. This handbook is intended to take the SQT developer through all the steps to consider when developing and validating an SQT. Sections are in sequence for normal SQT development.

1 -9. CHECKLIST. Most of the material in the handbook is neither difficult nor revolutionary. Developers will be familiar with some of it; other material will be new. To this end, the handbook is intended to serve as a self-administered checklist during development.

1-10. DIAGRAMS. The material is presented in the format of flow charts or decision diagrams. Handbook developers have attempted not to complicate your job with complex diagrams; therefore, do not let the diagrams turn you off without giving them a try.

Remember a few basic rules when "reading" the diagrams: an oval represents some sort of input or starting document or a final product; a rectangle represents some type of required action; and a diamond a decision, usually in the form of a yes/no answer. Other than that, simply follow the flow of the arrows.

1-11. TEXT. Diagrams are expanded on by text that explains how to do an action, or things to consider when you make a decision. As you become more familiar with the process in each section, you will rely less on the text and more on the diagrams. Even for experienced test developers, the diagrams serve as a checklist to insure that all aspects of SQT development are covered.

1-12. CAUTION. A word of caution is needed when using this handbook. The SQT program is a complex testing system that touches every job and skill level that exists in the Army Force structure.

No single document then, can possibly cover every question or situation that will arise in the various Test Development Activities (TDA) during test development. So while the handbook will serve as the basis for SQT development, there will be numerous exceptions when applied to individual development problems.

This handbook, then, applies to "normal" SQT development and is intended as an aid to development. The intent is not to have TDA circumvent the guidance when it suits them nor is it to hold them rigidly to unrealistic requirements.

1-13. GUIDANCE. This handbook also reflects TRADOC guidance. With any system as large and complex as SQT, further policy changes cannot be ruled out. Changes that affect development will be issued as changes to this handbook.

1-14. REFERENCES. There are some other publications you should read. These are:

- a. Letter HQ TRADOC Subject: EPMS Implementation Guidance and Milestones dtd 23 October 1975 w/ four incl. (The inclosures give you the technical and printing guidelines necessary for final production of your SQT.)
- b. "Handbook for the Administration of Skill Qualification Tests." Test Edition June 1976. (Contains guidelines for the Test Control Officer for SQT administration, but includes descriptions of materials to be provided by the TDA.)
- c. "Guidance to Proponents: Procedures to Prepare Manual for Administration" Letter ATTNG TMI-ITEG, HQ TRADOC, 2 April 1976.

### Section III. DEFINITIONS

1-15. COMBINATION PRODUCT/PROCESS TASKS. In HOC, the scorer may observe procedure for part of the task and observe outcomes or results for other parts of the task.

1-16. CRITICAL COMBAT SYSTEMS. The systems which are directly involved in the role of land warfare are termed combat systems. Combat systems are deemed critical when their operational presence is pivotal to the success of combat. "Systems" is used in its broadest sense, in that a tank is a system as is personnel replacement to combat units. Examples of standard critical combat systems include all nuclear capable weapons, intelligence gathering activities and individual anti-armor capabilities.

1-17. HANDS-ON COMPONENT (HOC). The HOC is that part of the SQT where the ability of the soldier to perform critical tasks is tested on actual job equipment or simulators. The HOC may take up to four hours to administer to the individual soldier. Total administration time at a typical installation will usually be two days but may extend up to five days for an HOC requiring major equipment end items or if it is to be administered to MOS with large population. Up to 16 tasks may be tested in the HOC.

1-18. INDIVIDUAL TRAINING EVALUATION DIRECTORATE (ITED). ITED is the activity designated by the Commander, TRADOC, to develop, implement and administer the SQT system. When you write to the ITED, use the following address: Director, Individual Training Evaluation Directorate, ATTN: ATTSC-IT, Fort Eustis, VA 23604.

1-19. ITEM. The basic unit of a written SQT is the item. An item consists of written presentation of information, or a question which examines answer by selecting the correct response(s) from alternatives presented. There may be as many as 10 alternatives per item.

1-20. JOB EXPERT. A job expert is a subject matter expert who has had recent field experience with soldiers in the SQT being tested, and has supervised the task in question. Generally job experts are NCO's and are oftentimes available at the TDA.

Although the terms "subject matter expert" and "job expert" are often used interchangeably, the distinction is important for the developer. For example, the industry representative for an item of equipment may be an SME but the individual who used or supervised the use of that item in a unit is a job expert.

Job experts are more useful to the developer because their experience qualifies them beyond technical competence. The job expert is always an SME, but the SME is not necessarily a job expert.

1-21. KEY ELEMENTS. To insure that the most important part of the task is tested, key elements will be identified. Key elements are the most frequent causes of failure or those elements which have serious consequences of failure. Key elements are identified by subject matter experts with field experience.

1-22. NON-PERFORMERS. The opposite of performers. Again self-ratings are primary method of identification, but other means also may be used. Level of training and experience often can be used as blanket predictors, e.g., soldiers in entry-level training courses are non-performers on all but the most basic tasks. Non-performers do not have to be totally deficient on the task in question. Non-performers are usually easier to identify than performers but their role is no less important in the validation process.

1-23. PERFORMANCE-BASED TEST. In the WC, some tasks or parts of tasks are tested by asking questions about how the task is performed. The questions are based on the examinee's knowing how to perform the task correctly even though he does not have to do it to answer the question.

1-24. PERFORMANCE CERTIFICATION COMPONENT (PCC). Many tasks require use of manipulative or other motor skills by the soldier, but also may be expensive tasks in terms of time, equipment and other resources. For that reason, a PCC may be included in an SQT.

The PCC allows a supervisor to observe a soldier's performance during the 12 months prior to the test quarter and to rate the soldier's performance using the performance measures, conditions, and standards contained in the Soldier's Manual.

1-25. PERFORMANCE MEASURES. After refinement, task elements become performance measures in the HOC. They are those steps or behaviors that the scorer looks at to decide if the soldier is performing the task correctly.

1-26. PERFORMERS. During validation of the components, it is necessary to identify soldiers who can be predicted to successfully complete the task being tested. The primary means of identifying performers is by the use of self-ratings, but may also involve supervisor evaluating experience, training and other independent measures of overall proficiency.

1-27. PROCESS TASKS. In the HOC, some tasks are scored by observing the process the soldier goes through in task performance. Each step or element is scored as it is performed in a process task.

1-28. PRODUCT TASKS. The opposite of process tasks. By looking at a result or outcome, the scorer can infer whether the task was performed correctly.

1-29. SCORABLE UNIT (SU). The term used to describe the means of testing a task for the WC, HOC, and PCC. In the HOC and PCC, a SU consists of up to 20 performance measures. In the WC the SU can consist of up to 10 items.

When scoring the SQT, one point is given for each correct SU. For normal SQT development, a single SU is used to measure each task, but you may have more SU if the task is especially complex and important. An SQT can contain between 30 and 76 SU.

1-30. SKILL QUALIFICATION TEST (SQT). The SQT is a validated performance-oriented test of the soldier's ability to perform the critical tasks required by his Army job. The SQT is used to assess job competence for training feedback and personnel management purposes. Each SQT may have up to three components.

1-31. SOLDIER'S MANUAL (SM). The SM is a book that lists for the soldier critical tasks needed to perform satisfactorily at his present skill level and the next higher skill level within his MOS. In addition, it tells the soldier how to perform the tasks, the expected conditions under which they will be performed, and the standards which must be met. The SM is the basis for the tasks used in the SQT.

1-32. SUBJECT MATTER EXPERT (SME). The SQT developer will be preparing tests for a number and variety of tasks. The developers' depth of knowledge of these tasks will vary. During development it is necessary to make use of individuals who have a thorough working knowledge of the task to assist in review of task analysis and answer questions during development. These individuals are generally available within the TDA, either as co-workers or within other departments.

1-33. SQT NOTICE. Many tasks cannot be tested in the SQT the same way they are detailed in the SM. The purpose of the SQT Notice is to tell the soldier which tasks will be in the SQT, and to explain to him how they will be tested. The SQT Notice is given to the soldier 60 days before the test quarter, and serves as the basis for training on those tasks contained in the SQT.

1-34. TASK. For SQT development only, a task is an action performed by a single individual which produces a result or product that can be measured (scored). A task has an identifiable start point and an identifiable completion point and intermediate actions that can be specifically stated in terms of behavioral activities.

(CAUTION: There are many different ways to define a task. This definition is not meant to be all encompassing or to serve other purposes which require definition of a task. It is intended only as a working definition for purposes of SQT development.)

1-35. TASK ANALYSIS. Task analysis is the written description of how a task is performed. It includes the following: conditions of performance, equipment and tools required, actions performed (broken out into steps or activities) and the standard for final products, results or outcomes. The task analysis is the basis for all developmental work in the SQT.

1-36. TASK ELEMENTS. The behaviors or steps in a task comprise the elements of the task. Elements are taken initially from the task analysis. It is sometimes necessary to add task elements prior to test development.

1-37. TEST CONTROL OFFICER (TCO). The individual responsible for the administrative functioning of SQT in the field. He is the point of contact between the field and ITED. He may be assigned within the operations staff section or an administrative staff section.

1-38. TEST DEVELOPMENT ACTIVITY (TDA). The TDA are schools and other organizations that develop and validate SQT and interact with the ITED in their review. There are approximately 34 TDA in the SQT program.

1-39. TRACKED TEST. Some duties and tasks within MOS may vary greatly in the way they are performed. This may be due either to differences in equipment, e.g., the M60A1 tank vs. M60A2, or because the jobs (duty positions) at the same skill level in an MOS are actually different, e.g., Mortar Crewman vs. Forward Observer in Skill Level 1 of MOS 11C. It would be unfair to test all incumbents within the MOS the same way. For this reason, some SQT have portions of the test which are tailored to a specific duty position or type of equipment. These are called tracked tests and are identified as such in the SQT Notice.

1-40. VALIDATION. An important aspect of the SQT is that all scorable units are pre-tested to assure that they discriminate between performers and non-performers and that they are perceived as fair by a representative sample of soldiers who will take the test. Validation also confirms that the items can be understood by the soldier and that HOC items are scored the same by different scorers.

1-41. WRITTEN COMPONENT (WC). The WC tests 30-50 tasks (scorable units). Each scorable unit is comprised of up to 10 questions, with each question having 2 to 10 possible responses. The WC is used to test those tasks requiring decisionmaking or other mental skills on the part of the soldier.

1-42. WRITTEN PERFORMANCE TEST. In the WC, some tasks or parts of tasks can be performed almost exactly like they are on the job. The tasks are characterized by the fact that the correct answer cannot be recognized without going through the task procedure.

ARI P-77-5

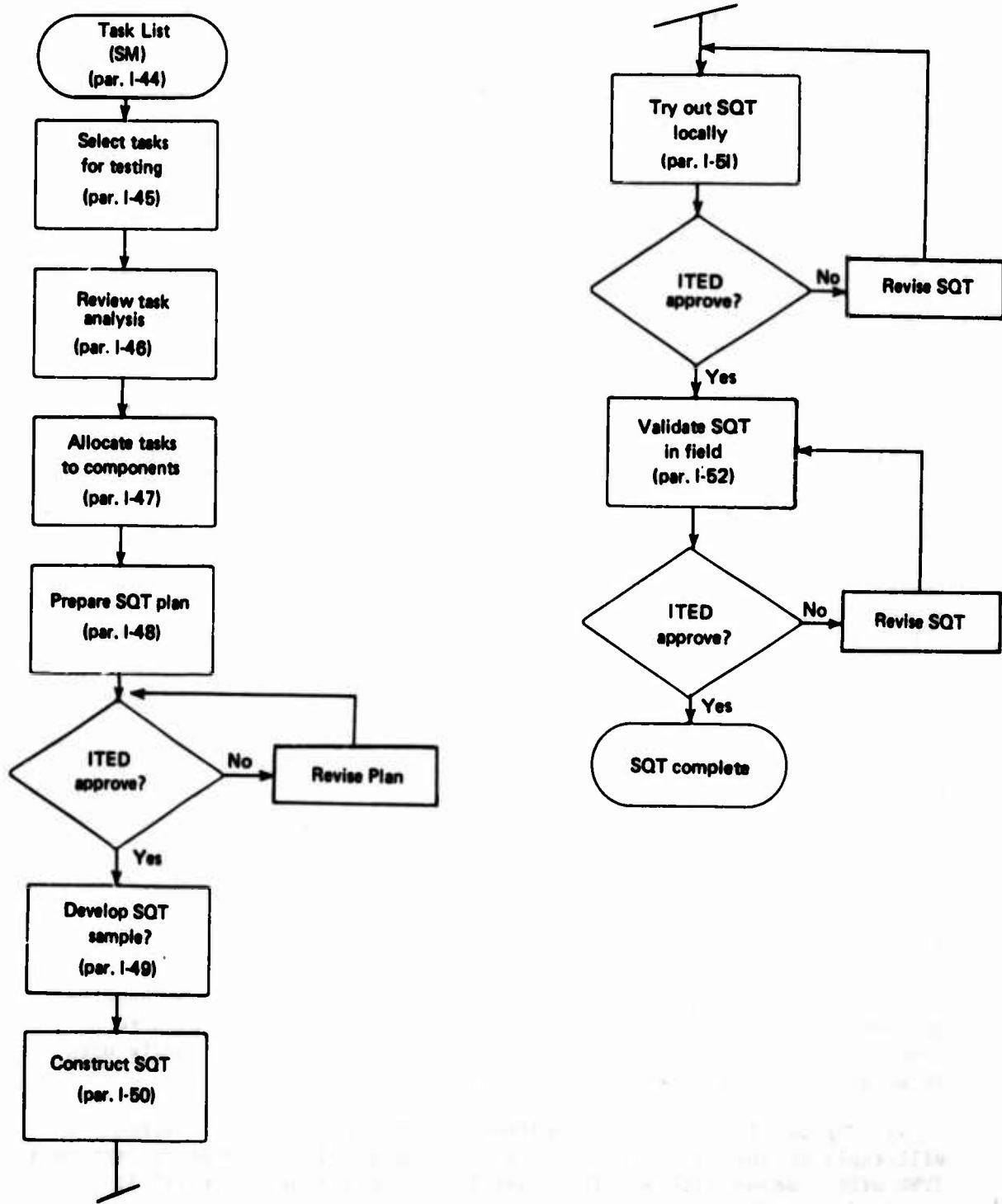


Figure 1-1. Overview of SQT Development Process

#### Section IV. SQT DEVELOPMENT PROCESS OVERVIEW

1-43. INTRODUCTION. The SQT development process includes constructing and validating each component. The process requires close coordination between the TDA and ITED. The development process is also dynamic because there are many opportunities to revise scorable units before the SQT is final.

1-44. SOLDIER'S MANUAL TASK LIST. The starting point for SQT development is the current or draft Soldier's Manual. Just as the Soldier's Manual tells the soldier what he must do, it tells you what group of tasks to test.

1-45. SELECT TASKS FOR TESTING. Most MOS have so many tasks that it is not practical to test all of them in an SQT. But selection of tasks is not random or haphazard. Tasks to test are determined through a systematic evaluation of field data and input from the senior officer responsible for the MOS.

1-46. REVIEW TASK ANALYSIS. Probably the single most important aspect of the SQT development process is review of the task analysis. The effort during the review process will determine adequacy of the final test.

1-47. ALLOCATE TASKS TO COMPONENTS. The SQT allows the developer three options for testing - the Written Component (WC), Hands-On Component (HOC), and Performance Certification Component (PCC). Currently the only requirement is that each test must have a WC. But there are desirable options of having an HOC and a PCC. Consider the characteristics of the task when you decide the best component to use for testing each task.

#### 1-48. PREPARE SQT PLAN.

a. You will prepare a comprehensive plan for each test you develop. The SQT Plan is submitted in two parts. One is a list of tasks to be tested by component; the other is an explanation of the rationale used in selecting and allocating the tasks.

b. The SQT Plan must be submitted to ITED for approval. ITED will reply to the plan within two working days. If the plan is deficient, ITED will respond with specific questions. Once final approval is obtained from ITED, the plan becomes the guide for all developmental work in the SQT and the plan cannot be changed without ITED approval.

1-49. DEVELOP SQT SAMPLE.

a. When the SQT Plan has received local (TDA) approval, work can begin on the SQT sample. The following items must be submitted to ITED as part of the sample:

- (1) Five written scorable units.
- (2) Two hands-on scorable units (if used).
- (3) All performance certification scorable units (if used).
- (4) The SQT Notice (less the component lead-in paragraphs) for the scorable units submitted in the sample.
- (5) The task analysis for the scorable units submitted in the sample with key elements identified.

b. The scorable units must be submitted as complete as they will be for testing. ITED will reply to the sample within 10 working days after receipt. If the sample is approved, work may begin on development of the complete test. If the sample is not acceptable, ITED will return specific guidance for corrections. You must make the required changes and submit a second sample for any deficient component.

1-50. CONSTRUCT SQT. As soon as ITED approves the sample, begin developing the rest of the SQT.

1-51. TRY OUT SQT LOCALLY. You must try out each hands-on and written scorable unit locally before you submit the proposed SQT to ITED. This tryout is the first step in validation. Check out each scorable unit for completeness, job relevance, doctrinal accuracy and perceived fairness. Also check the clarity and readability of WC items and the reliability of scoring for the HOC.

1-52. VALIDATE SQT IN FIELD.

a. The field validation for the HOC has two phases. The first phase is to check the reliability of scoring each scorable unit when representative soldiers are tested. When this is done, the TDA, in conjunction with ITED, selects a field unit for tryout of the complete HOC.

The TDA submits the entire component including administrative instructions to the Test Control Officer for the unit selected. The chief purpose of this validation run is to check the administerability of the HOC. The TDA validates the WC on a group of up to 30 soldiers, divided approximately equally as performers and non-performers.

The objective here is to establish the discriminant validity of each written Scorable Unit. The TDA also has the PCC validated subjectively at this stage. The PCC is submitted to a sample of commanders and supervisors who assess its relevance and administrative feasibility.

b. Results of the field validation are used by the TDA in preparing a final draft of the SQT. When approved by ITED, the SQT is ready for operational administration which concludes the development process. Printing and distribution, and later, scoring, data reduction, and distribution of test results are all carried out by ITED.

## CHAPTER 2

## SELECT TASKS FOR TESTING

## 2-1. INTRODUCTION.

a. Each skill level of an MOS is defined by a list of tasks that a soldier is expected to be able to perform. The comprehensive task list is contained in the appropriate Soldier's Manual. The SQT will include tasks from a soldier's present and next higher skill levels.

As a general rule for skill levels one through four, two-thirds to three-fourths of the tasks in the SQT will measure the soldier's competence at his present skill level, while the remainder of the one-fourth to one-third of the tasks will measure his competence at the next higher skill level.

For the higher skill levels or when MOS merge, the mixture may be up to 50/50, and in some MOS there may be little or no distinction among tasks between skill levels.

b. The test developer is faced with the problem of selecting a small number of tasks from all the tasks in the soldier's job. The SQT is not necessarily intended to be a random sample of the different kinds of tasks that a soldier is expected to be able to perform.

This distinction is important. Tasks are not selected for testing merely because they are representative or because most soldiers in the MOS must be able to perform them. Testing resources are a precious commodity and tasks should be selected that will provide a high "payoff" for the effort necessary.

To a large extent, SQT testing will influence training so the test developer has a further responsibility in his task selection. This section is concerned with the procedure for selecting "best" tasks for inclusion in the SQT.

c. Before discussing methodology for task selection, it is important to emphasize the role of the senior officer in task selection. The senior officer may be the Assistant Commandant (AC) or the Director of Training Development (DOTD) at the TDA.

In either case, it is important to obtain the senior officer's guidance and concurrence in task selection. You, as the test developer, must insure that the chain of command is exercised early in the task development process to insure concurrence with the method of task selection and with the tasks selected.

The senior officer is a valuable resource for identifying tasks for testing because he has the broadest perspective. The procedure will certainly vary at different TDA. The requirement to get the senior officer involved in task selection does not imply that each discrete task must have General Officer or O-6 approval.

Categories of tasks and the method of selection must reflect the senior officer's input before development is started.

d. No single method of task selection will meet the requirements of all TDA. We will discuss them here without regard to priority, although some methods are more desirable. The TDA must select the method or combination of methods that will provide them with the best sample of tasks based on information available. Most TDA will use more than one method. The important thing to ask is, "Will this task selection method provide the highest payoff for the resources I'm expending on this test?"

e. The methods are:

- Known performance deficiencies.
- Critical combat systems.
- ARTEP's.
- Evaluation results.
- Job content.

2-2. KNOWN PERFORMANCE DEFICIENCIES. The purpose of this method is to select tasks that tend to be performed poorly by most people in the MOS. If only one method of task selection could be used, it would likely be this one. The difficulty with its use is in obtaining sufficient information on which to base judgments. Below are discussed some more generalized sources that should be considered.

a. TOE Manpower Authorization Criteria (MACRIT). DARCOM maintains a master data file of annual direct productive maintenance man-hours (AMMH) for all standard and limited procurement items of equipment (as listed in AMC Supplement One to AR 570-2, Manpower and Equipment Control).

This file is maintained at USAMMC and the AMMH data is updated annually by the commodity commands. The AMMH data is furnished to TRADOC requesting activities. The AMMH printout includes operator/crew maintenance requirements codes, supporting maintenance MOS and Organizational, DS and GS AMMH breakouts.

b. Red Team Assessment Programs. The Red Team Assessment program is designed to provide CG, DARCOM with an assessment of Selected Critical Field Systems. The appendixes to this circular list the systems and the schedule of assessment reporting.

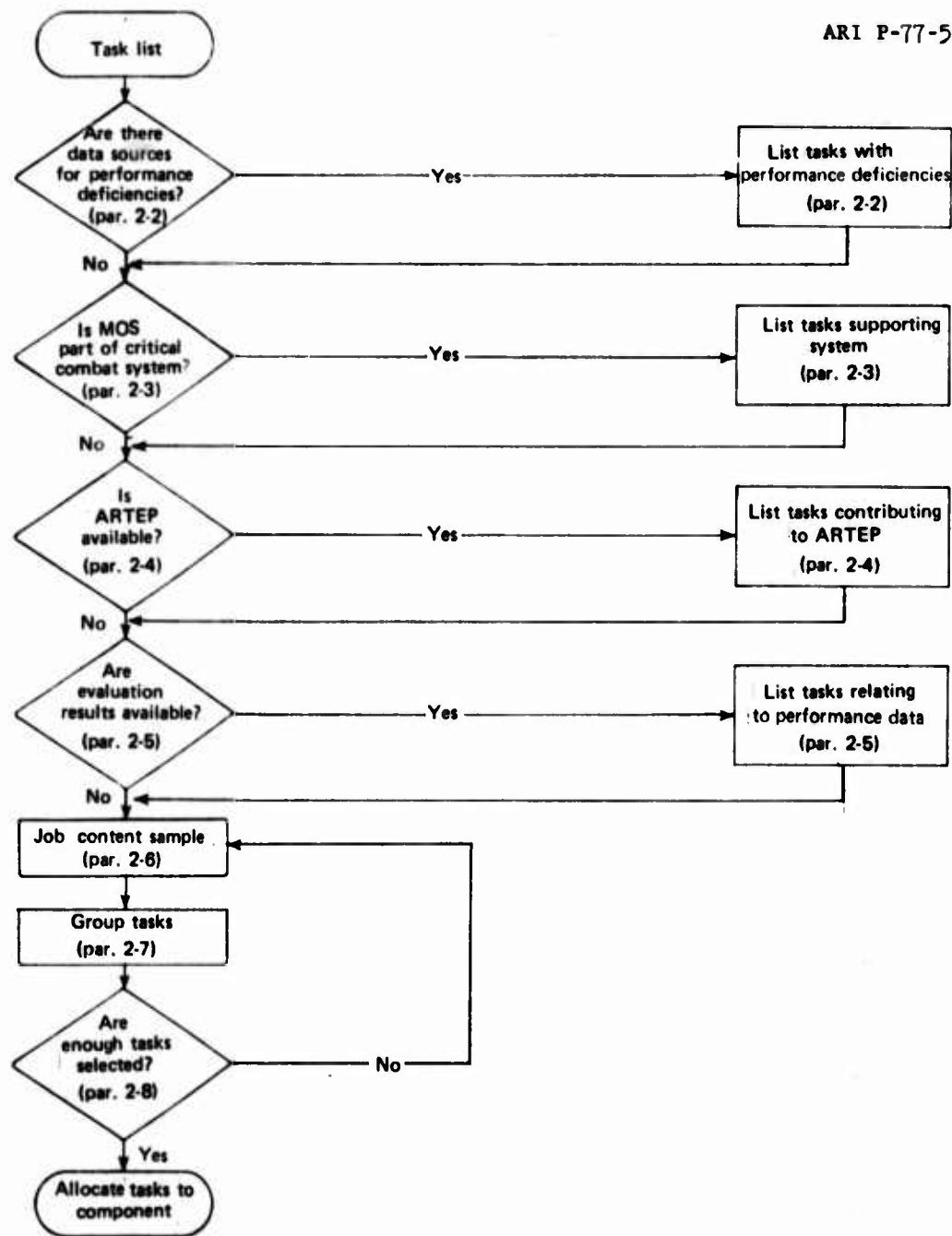


Figure 2-1. Select Tasks for Testing

c. U.S. Army Maintenance Management Center (AMMC) (Bluegrass Army Depot, Lexington, KY 40511) collects and maintains a variety of equipment-maintenance related reports and data. These include:

(1) DA Sample Data Collection (SDC). AR 750-37 specifies a procedure for collection of maintenance data utilizing a sampling technique. The U.S. Army Maintenance Management collects and analyzes data on selected items of equipment or to study specific problems.

(2) Reports furnished by commodity commands to AMMC for Command specific items and problems. These reports identify equipment with high degrees of fault and deficiencies.

d. Technical Proficiency Inspections (TPI) and Technical Standardization Inspection (TSI). These results are available from major Headquarters such as HQ USAREUR and 7th Army and FORCE command.

e. IG inspection results are also available from Major Command Headquarters. They should be requested by specific topic.

f. Direct Mail Questionnaires (DMQ).

(1) For some TDA, the only method of identifying specific deficiencies may be through the use of a direct mail questionnaire to field commanders. This may be used also in conjunction with any of the sources above. The level of command to obtain the information from is best determined by the TDA.

Too low a level may reflect only a local view (plus increase the number of responses which must be tabulated) while too high a level in the command structure may not accurately reflect the specific deficiencies.

The format of the questionnaire can be tailored to fit the type of information desired by the TDA, but the information requested should be as specific as possible. Do not ask the commander for a subjective analysis of what he thinks should be tested.

(2) CAUTION: Before using the DMQ, you must coordinate with MILPERCEN. Data banks already exist on many of the CMF and the information you seek may already be available. MILPERCEN must also approve questionnaires sent to the field. For MILPERCEN guidance on use of mail questionnaires coordinate through ITED.

g. It is recognized that most of the above sources (except the DMQ) are hardware or equipment oriented and may not be equally useful to all TDA. But each TDA should evaluate information available from these sources before discarding them. You must also be aware of the time requirements involved in using these sources.

Although the sources are responsive in providing available data, it must still be analyzed. A test developer who faces a two-week deadline to ITED is probably not in a position to access and analyze the information. **TAP THE RESOURCES EARLY**, preferably by maintaining updated source documents.

### 2-3. CRITICAL COMBAT SYSTEMS.

a. Tasks chosen for testing may be selected on the basis of their contributions to the successful operation or maintenance of critical combat systems. Critical combat systems are equipment items or other operations that are pivotal to a unit's success in combat.

The critical combat system will vary from one MOS to another as will its usefulness as a task selection tool. Some TDA will tend to feel that almost all of their tasks are related to critical combat systems, while other TDA may find their links to critical combat systems tenuous at best.

For either, the critical combat systems approach will be of limited use. The usefulness of the approach occurs primarily in the area of combat service support and to a lesser extent in combat support where a distinction can be made regarding activities relating directly to critical combat systems.

For example, the tasks that are directly involved in divisional support of maneuver units would be more suitable for selection than tasks which are not. An example of the critical combat system approach occurs in the track vehicle mechanic MOS supporting an artillery unit.

The tracked carrier for an artillery piece having nuclear capability is a critical combat system. Tasks directly related to repair of that vehicle would be more important for testing than tasks required to repair a track vehicle without a nuclear capable piece.

b. To use the approach, first identify the critical combat systems that are related to the MOS, then list, from the Soldier's Manual, individual tasks that relate directly to that system.

2-4. ARTEP.

a. The Army Training Evaluation Program (ARTEP) should be examined when considering tasks for SQT development. The ARTEP, of course, describes unit tasks while the SQT is based on individual tasks. Therefore, you must identify the individual tasks comprising unit tasks or missions in the ARTEP.

b. Presently, consolidated ARTEP results are not available, however individual unit ARTEP experience may be gathered through the DMQ described above. Validation results of ARTEP are available, however (usually in TDA Directorate of Evaluation) and these should be reviewed. Here again, look for deficiencies in crew-unit performance that can be translated into individual tasks for use in the SQT.

2-5. DIRECTORATE OF EVALUATION (DOE). Most TDA already have established a Directorate of Evaluation and this organization has the responsibility for varied evaluation activities proponent to the TDA. You must be familiar with the information gathered by the DOE and match the information with the SQT development

Besides the above mentioned ARTEP validations results, the DOE will most likely be the repository of results from SQT already administered.

2-6. JOB CONTENT. Just as known performance deficiencies are the most desirable tool for task selection, sampling on the basis of job content is probably least desirable. But it may occur within a TDA that no alternative exists or that even if above sources are used, too few tasks have been selected. If the expertise and experience available to you is insufficient to enable you to make judgments as to degrees of criticality for testing, you will probably need to resort to sampling from the Soldier's Manual to "fill out" the SQT. However, even here the sampling is not done on a totally random basis. Tasks should be organized into content or functional areas, e.g., duty categories, and a proportional sample drawn from each. This provides a combined sample representative of the full range of tasks in the Soldier's Manual. The resulting sample of tasks should be approved by senior officers as discussed before.

2-7. GROUP TASKS. If you used more than one method or source for task selection, you may have an unorganized list of tasks for scorable unit development at this point. If so, sort the tasks in order of priority for testing. This need not be a precise rank ordering.

Tasks may be roughly grouped, e.g., top, middle, bottom, in terms of priority for testing. The order of this list is not final and no tasks are being discarded at this point. Ordering of tasks is only a tool to use in planning allocation of development efforts.

The purpose of this grouping is two-fold. First, you will use it when allocating tasks for testing in Chapter IV. Just as important; you may have identified tasks that you will want to test regardless of subsequent testing restrictions that you may encounter. For example, your AC may have specified tasks that must be tested.

You certainly want to identify the tasks in some manner as having the highest priority for testing. You may have other tasks that were identified by less reliable criteria which you may consider dropping later from the test if problems develop with size or feasibility.

In ordering the tasks, developers may use differing rationale, but base your list on the information received from the sources listed above. The larger the list of tasks from these sources, the more important the ordering process.

2-8. HAVE ENOUGH TASKS BEEN SELECTED? At this point you should have a rough idea of how tasks will logically fit together into scorable units. You should have a pool of about 30 tasks at this point, even though you have not decided which component is best for testing any of the tasks. If you do not have enough tasks, add tasks, always using data available from one or more of the task selection methods.

2-9. ORDERING. You should have now a roughly ordered list of tasks that, in addition to the basic data sources, reflect guidance from senior officers. Now, a final review by the AC or Director of Training Development should be obtained for concurrence prior to actual test development.

2-10. TASK SELECTION. Although the need to start the task selection process early has been stated before, it bears reemphasis. In the past, since some developers have been faced with the problem of short lead time to meet suspense dates for SQT submission, task selection has been based on a limited and subjective analysis.

To insure that task selection is based on the best information available, the selection process must be started well in advance of the development phase. It is much easier to make adjustments with later information than to force a hurried and inadequate task selection.

CHAPTER 3

REVIEW TASK ANALYSIS

3-1. INTRODUCTION.

a. Importance of a good task analysis in SQT development cannot be overstated. Experience has shown that when test developers have problems in later stages of development, the trouble can be traced directly to deficiencies in the task analysis.

A thorough task analysis is essential to the development of a good scorable unit, that is, a scorable unit relevant to measurement of task performance. What the developer does between task selection and final validation is based on presumed existence of a good task analysis.

A "bad" task analysis will result in an invalid test - a test that simply does not measure what the task requires. No matter how well you do your job, if the work is based on an inadequate task analysis, the test will not be good.

b. Task analysis probably still does not follow a standardized format among TDA and may be referred to as a task data card, training analysis worksheet, task detailing or task structure analysis.

The name or format is unimportant for the developer although it must be uniform within the TDA. TDA will also vary in terms of how active a role SQT developers play in task analysis.

However, it is important that SQT developers have input into the task analysis program. Task analysis serves many purposes, one of the most important being SQT development. Therefore, you, the developer should make known your needs to whoever is responsible for task analysis.

During review of task analyses, gaps and errors are often discovered. It has been the experience of some TDA that once a task analysis has been approved, it is extremely difficult to change. This tendency must be overcome. The task analysis is the building block from which both the SQT and the Soldier's Manual are constructed; all three must track in the final review.

c. It is not the intent here to teach you how to do a task analysis (that would be another complete area of study). Nor is it intended to criticize existing or ongoing task analysis efforts.

Rather it is intended to communicate how to review a task analysis to insure it contains information needed for test development. It is important that you follow a logical review process in order to identify, as early as possible, information gaps in existing task analysis data.

When deficiencies are discovered, you most likely will have to use your own resources to gather required information. But you must also insure that you coordinate with the people responsible for task analysis to insure they are aware of your problems and requirements.

3-2. ARE TASK ELEMENTS THE SAME IF CONDITIONS VARY?

a. Task elements are steps or procedures in performing the task. Conditions are factors which influence performance of the task. Most conditions are associated with environment, that is, weather, temperature, light, precipitation and the like. Conditions may also include equipment, terrain, mission or location.

b. The task analysis generally will contain a statement of conditions under which the task is performed. Some tasks may change in how they are performed if conditions are changed. Therefore the developers must consider the task under all probable conditions.

Sometimes the "changed" task will become a new, separate task statement. Changing conditions sometimes alter steps or elements in task performance, and sometimes simply make the task more or less difficult.

For example, disassembly of the M16 rifle in the dark under freezing rain, is certainly more difficult than indoors at a well-lighted work-bench. But the elements in performing the task remain the same. Other tasks, however, will be changed by conditions, e.g., vehicle maintenance after fording operations.

3-3. IS THE TASK CRITICAL UNDER VARIOUS CONDITIONS?

a. If you decide that a task has different elements under different conditions (and is not included in the existing Soldier's Manual or task analysis), you may have to decide if the "new" task is critical for testing.

But before that decision can be made, those responsible for the Soldier's Manual must first decide if the task warrants inclusion in the manual. If the task is not judged to be job-critical, and therefore not in the SM, it does not become a candidate for testing.

On the other hand, if task analysis personnel decide the new task belongs in the SM, you must then decide whether to include the task in the SQT. Follow the procedures outlined in Chapter 2 when you make that decision.

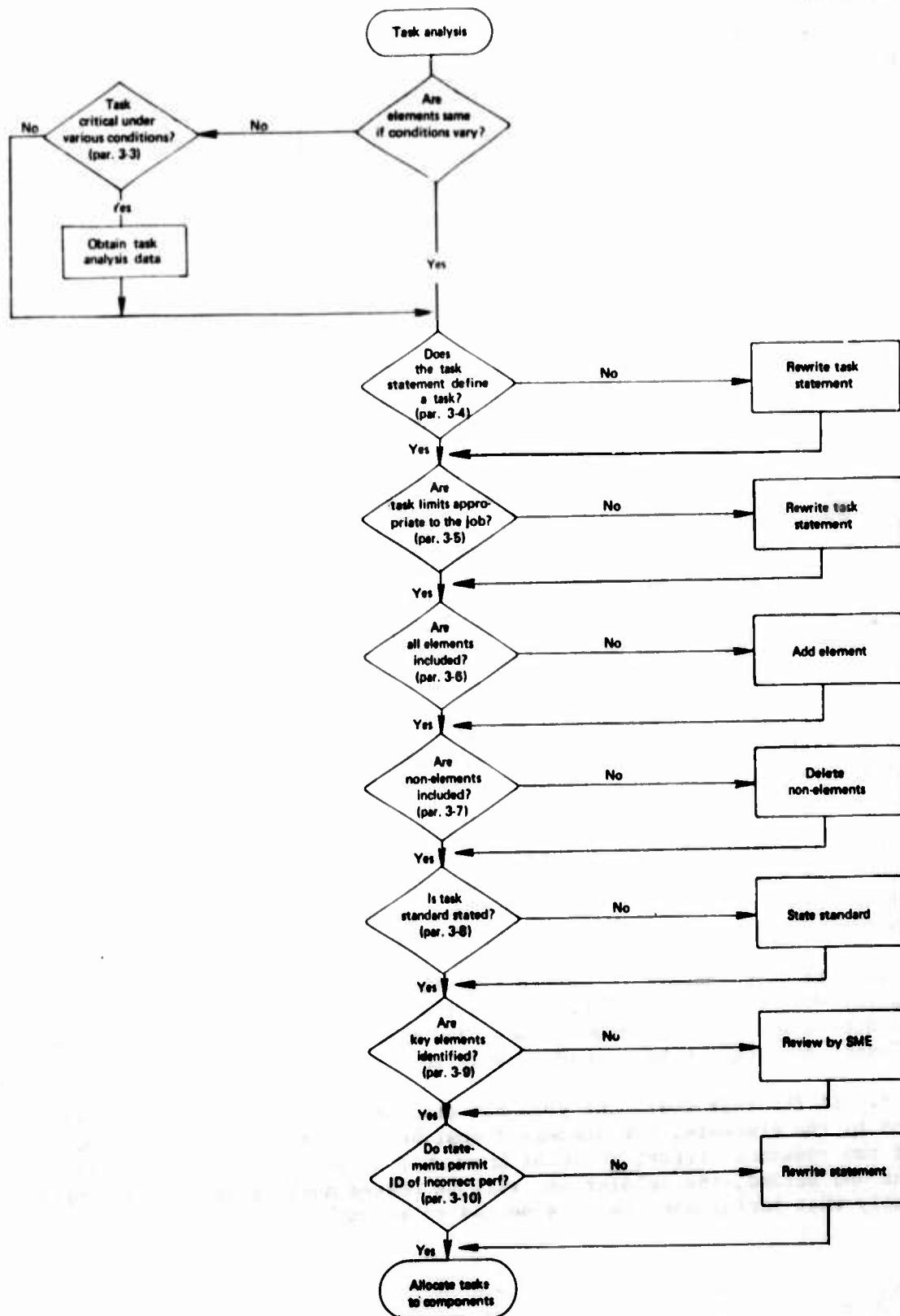


Figure 3-1. Review Task Analysis

b. If elements of the task remain the same under various conditions, there is no requirement to obtain more task analysis data. If the elements do change, treat the task like a new task.

- (1) Is the new task critical? If the answer is "no," discard the task.
- (2) If the answer is "yes," obtain task analysis data and review it.

#### 3-4. DOES TASK STATEMENT DEFINE A TASK?

a. Definition of a task is imprecise and although a quantity of literature exists on defining a task, again it is intended not to involve the test developer in an area that belongs to task analysis personnel.

However, during review of the task analysis, insure that the task statement describes what is to be performed. The task statement must define an action and the task statement must have a measurable outcome.

For purposes of SQT, knowledges are not considered tasks; the essence of a job task is doing something, not knowing something. Compare the task elements as listed with the task statement and insure that the statement is the activity described.

Be cautious about tasks that are actually sub-tasks. Some sub-tasks may be tested as tasks, depending on the skill level of the MOS involved. But normally the end task should be tested, not a series of sub-tasks tested separately.

b. Be careful of generalized statements expressing attitudes, methods or non-concrete behaviors. Generalizations often arise when attempting to describe duties of NCO's. Be alert for verbs such as "inspect," "supervise," "direct," "counsel."

It will be necessary to obtain more task analysis information as to exactly what these incumbents do. If the actions are not identified at the start, you may well expend much time and effort only to find you have a "task" that cannot be tested as written.

c. If the task statement does not describe a task or does not correspond to the elements, the statement must be rewritten. This is important for two reasons. First, you must know what activity you are going to test and second, the soldier who will be tested must be able to identify easily what performance he is expected to accomplish.

d. You may find a task statement too broad for what should be tested at a certain skill level. Yet because the Soldier's Manual has been distributed, you find the task analysis is "locked-in." If the broad task is actually an aggregate of several subtasks, you should treat the subtasks separately for test purposes and use the SQT Notice to inform the soldier of additional "tasks" to be tested.

### 3-5. ARE TASK LIMITS APPROPRIATE TO THE SKILL LEVEL?

a. This decision is similar to the preceding decision but focuses on a more precise aspect of soldiers' performance. Task limits are parameters of the job as they apply to soldiers at the skill level considered. The developer is deciding where the task, as performed by the soldier, begins and ends and what activity it encompasses.

b. Task limits means different activities for different levels of soldiers. A first sergeant may prepare a draft reply to a congressional inquiry and his clerk may type it. Yet the task "Prepare Correspondence" is often listed for both duty positions.

c. Problems with task limits are caused generally because the task statement is too broad. Restate the task in terms of what the incumbent performs, then check to insure that task elements accurately describe that performance.

### 3-6. ARE ALL TASK ELEMENTS INCLUDED? Task elements are component behaviors required to perform the task. Elements roughly correspond to steps, behaviors, or actions listed in the task analysis. Ask the following questions which may add to or modify the elements, depending on the task.

a. What initiates the task? If the task is self-initiated, consider including the initiating cue as an element.

b. What completes the task? If the end of the task is a condition or product requiring judgment on the part of the incumbent, consider including the end product as a stated element.

c. Is use of references necessary? If a reference (TM, FM or other job performance aid) is used normally on the job, that fact should be indicated in the task analysis. References necessary for task performance on the job generally will have to be available for test purposes.

This availability poses no problem for hands-on testing, but may pose a logistical and cost problem for the Written Component which must be self-contained. To minimize the problem ask, "Is the ability to use the reference, or the ability to use the information in the reference, the important part of the task?"

If ability to use a manual is important, this should be treated as a separate scorable unit within the hands-on component. If ability to use information is the important part, extract and include the information with the relevant scorable units in the Written Component. Coordination with ITED is necessary for any combination of extracts exceeding 30 pages per SQT.

d. Are all elements stated? Insure that a concise statement for each step is in the task and that no actions have been left out. You may not use all the steps when you develop performance measures, but it is necessary to start with all of them.

3-7. ARE NON-ELEMENTS INCLUDED?

a. Once you are sure all elements have been included in the task analysis, check to see if any unnecessary elements are listed. Delete from the task analysis nonessential elements. For example, if the task analysis data on target engagement includes knowledge of the muzzle velocity of the bullet, that knowledge should be deleted as nonessential to performance of that task.

b. Finally, mark key elements on the task analysis data form by asterisks or other means. These are the elements on which to focus attention when developing scorable units.

3-8. IS THE TASK STANDARD STATED?

a. The task analysis should include a statement as to standard of performance. Standards state how well a task must be performed and often highlight unacceptable actions. Standards tell how to judge the adequacy of performance and may be included as an element in the analysis or listed separately. Always consider possible standards of sequence and time.

b. Is sequence important? If the task has a required sequence, then the sequence must be stated. Only those aspects of sequence absolutely necessary for correct performance or safety to soldier or equipment must be specified. Be alert for tasks where only some steps must be done in sequence. Be specific when describing sequence. For example, "All steps must be performed in sequence," or "Step five must be performed last."

c. Is the task time constrained? If the task is time-constrained as a reflection of job conditions, the time standard should be specified. Do not confuse time-constrained tasks with time limits routinely stated in the SM, or that may be adopted later for test administration. Examples of true time constraints are the nine seconds to put on the protective mask, or two minutes to remove a misfire.

3-9. ARE KEY ELEMENTS IDENTIFIED?

- a. Use of key elements is unique to SQT development. So it is unlikely that key elements have been identified during task analysis.
- b. Key elements are those elements which have been identified as:
  - (1) Being most common sources of failure in performance of the task.
  - (2) Having serious consequences of failure.
- c. The first aspect of key elements answers the question, "Why do most soldiers fail to perform this task correctly?" The aspect is based on frequency or the "most likely" step where the soldier will fail. The aspect may be based on data gathered during evaluation of task criticality (Chap. 2) or on collection of opinions from subject matter experts.
- d. The second aspect of key elements pertains to serious consequences of failure. Many, but not all, elements in this category are related to personnel or equipment safety. The elements are not related to frequency; in fact, they may occur very infrequently.

But consequences of their occurring are so serious as to warrant testing them.

You must be somewhat selective in using this category of key elements. Certainly not all safety warnings qualify for inclusion. Information on this can also be gathered by interviewing subject matter experts.

e. Not every task will have key elements, but do make every effort to identify them. Developing the SQT around key elements will help insure that testing efforts result in the greatest payoff.

f. At the end of this chapter is a sample questionnaire that may be used in gathering data from subject matter experts about task elements and key elements (Figure 3-2). No magic formula exists for analyzing the data on key elements once it is collected.

Look instead for trends or consistencies in replies. Most experts will identify more than one element or step as being the most frequent reason for failure. If agreement among experts is good, you should have no problem in identifying the key elements.

On the other hand, if agreement is poor, you may have to follow-up by discussing sources of disagreement with experts in attempting resolution. Bear in mind, however, the possibility that no one or two elements can be pinpointed as common sources of failure.

g. Significance of key elements will be obvious when you develop scorable units, particularly written scorable units. When tasks have many elements, concentrate on preparing items that focus on key elements.

Additionally, several written items likely will be developed around key elements to insure at least one item per key element survives validation.

**3-10. DO STATEMENTS PERMIT IDENTIFICATION OF INCORRECT PERFORMANCE?**

a. After key elements have been identified, examine each element statement to determine whether an adequate basis exists for identifying specific errors common in incorrect performance. Often the statements are too general to be of much use.

Ask yourself, "What does this statement mean?" and, "Can I tell from this statement if a soldier is correctly or incorrectly performing this step?" Add to the statement, if necessary, a standard or way to measure the performance.

Compare, for example, "adjusts volume" with "turns volume knob to LOUD position." Compare "cleans metal surfaces" with "removes all traces of dirt and rust, using solvent and #2 emory paper. Wipes dry with clean soft cloth. Applies light coat of PLL special and again wipes surface dry."

b. Statements that do not allow identification of incorrect performance must be revised. While TM, FM, and other publications will provide you much needed information, it may be necessary to seek subject matter expertise. Insure statement accuracy by checking with several job and subject matter experts.

Sample Questionnaire

IDENTIFYING KEY ELEMENTS IN  
TASK PERFORMANCE: QUESTIONS FOR A JOB EXPERT

1. Is the task analysis a complete listing of all elements (decisions or actions) necessary for successful task performance?

Yes \_\_\_\_\_ No \_\_\_\_\_

If not, what elements should be

a. Added? Please List:

---

---

b. Deleted? Please list:

---

---

c. Revised? Please list:

---

---

2. You have seen this task performed many times in the field. Of the times you have seen soldiers fail to perform it successfully, why have they failed? Where do they usually go wrong? Please consider the original list of task elements, plus any revisions you may have made, and identify the element that is the most frequent cause of failure.

Most frequent cause \_\_\_\_\_

Are there other task elements that stand out in your memory as frequent causes of failure? If so, please list them.

---

---

3. Sometimes performance of a task involves an element that is very important in the sense that failure to perform that element can lead to serious consequences, serious injury to the soldier, or serious damage to equipment. Is such an element part of the task being reviewed?

Yes \_\_\_\_\_ No \_\_\_\_\_

If there is, what is it?

---

---

Figure 3-2. Sample Questionnaire

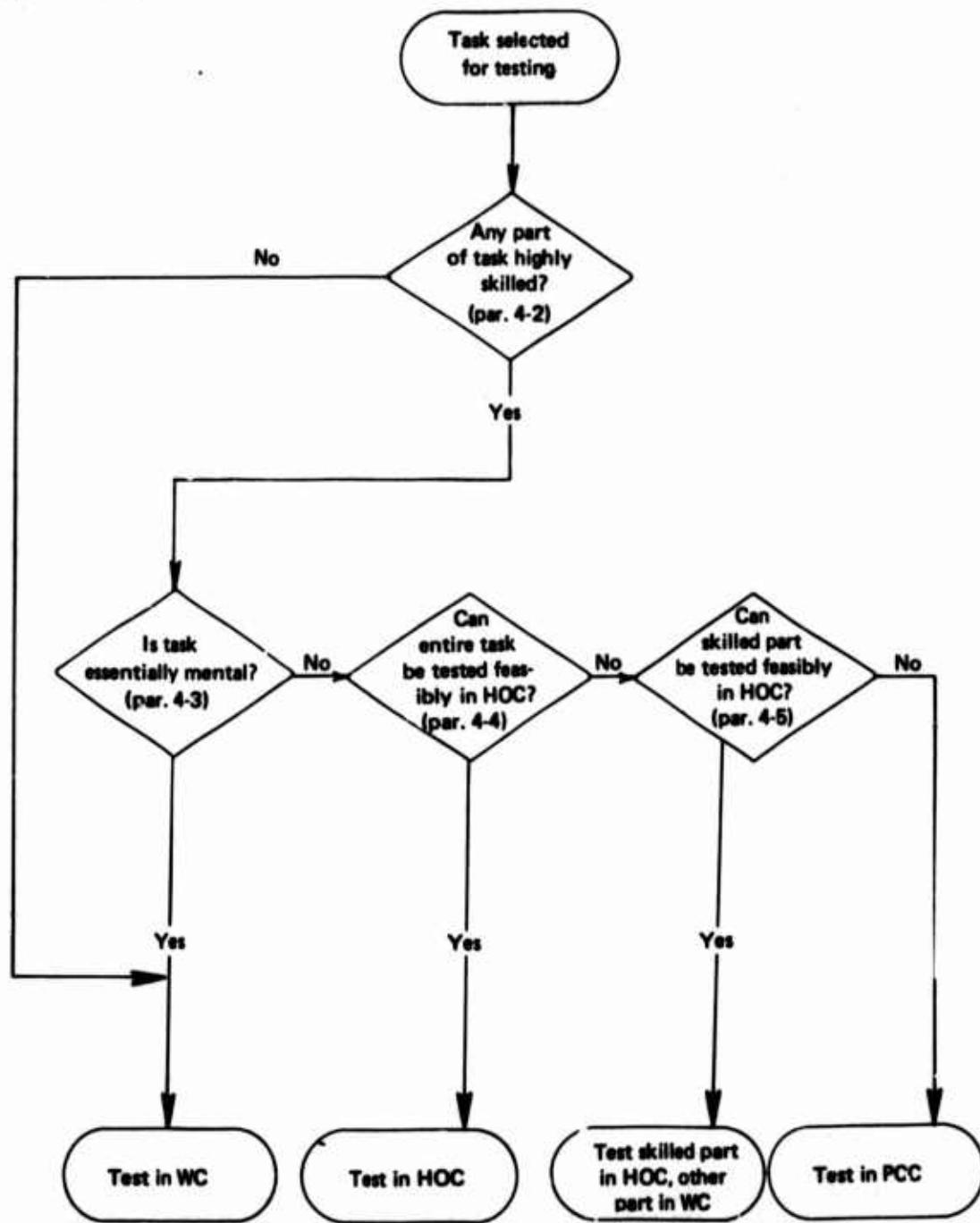


Figure 4-1. Allocate Tasks to SQT Components

## CHAPTER 4

### ALLOCATE TASKS TO SQT COMPONENTS

---

#### Section I. TASK ALLOCATION

##### 4-1. INTRODUCTION.

a. After tasks are selected for testing and review of tasks analysis information is completed, the next step is to allocate each task (or in some cases, part of a task) to one of three components (HOC, WC, and PCC) of the SQT. Basic information inputs to use in the allocation process are:

- (1) A list of tasks roughly ordered as to importance for testing.
- (2) A list of elements for each task with key elements identified.

b. Since tasks will be grouped with respect to importance or criticality, the general procedure is to start at the top of the list and go from "most" important to "least" important until the allocation process is complete. When there are a large number of tasks, this process will insure that the important tasks are included in the SQT.

c. Two basic concepts govern the process for allocating tasks to SQT components. Concepts are task content and testing feasibility. Both concepts assume that you know how the task being considered is performed.

If you have gone through requirements in Chapter 3, Reviewing Task Analysis, you will know how the task is performed, even though you may never have seen the task performed. If unsure about part of task performance, check with a subject matter expert. Keep in mind that concepts may be applied to different parts of the task.

##### (1) Task Content.

(a) This concept involves deciding on the type of task and the degree of skill required. Two general types of tasks are considered - mental and physical - and two levels of skill - high and low.

(b) Think of physical tasks as those that require use of hands, arms, legs, or body in performance. Think of mental tasks as those that require primarily use of the reasoning or thinking process in performance. Be aware that considerable overlap exists.

All physical tasks require some degree (often very high) of mental activity. All mental tasks require some physical activity, if only pushing a button or writing something. Be aware of the purpose of the classification of tasks by type - to decide what is needed for measurement.

Ask yourself, "If I watched a soldier perform this task, what would I observe?" If there is a lot of physical activity, the task is probably physical. If most of the essential activity cannot be observed (because it goes on in the soldier's head), the task is probably mental.

(c) Likewise, do not confuse skill with ability or level of the soldier. Skill (high or low) has to do with how much practice is needed to become proficient in the task--mental or physical.

If a soldier can read, or be told, about a task and then perform it with no more than one practice trial, the task is probably low skill. The more complex a task becomes, the higher the skill generally required.

(d) Certainly a degree of judgment is required in the above decisions, so do not be concerned with precise labeling of all tasks. Also, you may change your mind about tasks or subdivide parts of tasks into different classifications.

Be cautious about subdividing. Don't try to break a task into too many categories - you will end up with something probably untestable. Keep in mind the reason for the concept of task content. All types and levels of skill can be tested validly in a hands-on mode, assuming unlimited resources.

Some physical tasks may be tested in the written component, but highly skilled physical tasks must be tested in the hands-on component or performance certification component. Some mental tasks may need to be tested in the hands-on component, but mental tasks are usually more amenable to variations that permit written testing.

(2) Testing Feasibility.

(a) Testing requires some resources. In the case of SQT, the written component is the least demanding since, other than occasional use of job aids and AV media, it requires only paper and pencil. Moreover, many soldiers can be tested on many tasks in a relatively short time with one test administrator.

The hands-on component, however, requires equipment, terrain, and several qualified scorers. In addition, relatively few tasks can be tested in four hours.

The performance certification component was set up to accommodate tasks that are extremely time consuming and that require special terrain, facilities, or expensive equipment.

(b) With an understanding of the above concepts, take each task through the flow chart at the beginning of this section.

#### 4-2. IS ANY PART OF THE TASK HIGHLY SKILLED?

a. "Any part" refers to blocks of elements - not just single elements unless it applies to single elements identified as key elements (see Chapter 3 for identification of key elements).

b. "Highly skilled" should be thought of in terms of how much hands-on practice is required to become proficient at the task. Or, from the opposite point of view, can the task be performed by soldiers who know the procedure even though they actually may have done it only once or twice? Some useful clues to identifying high-skill physical tasks are:

- Task involves finely tuned acts of physical coordination, e.g., tracking a target, engaging targets with an M16.
- Task requires quick reaction; must be performed with tight time limit, e.g., clearing a jammed weapon, typing a message, donning a gas mask.
- Task involves time-sharing - doing two things at once, e.g., tactical driving.

c. If no parts of the task involve highly skilled actions, the task should be considered for the WC.

#### 4-3. IS THE TASK ESSENTIALLY MENTAL? If a part of the task is highly skilled, determine whether the skill is primarily mental or physical. Is the skilled part of the task a mental operation or a series of mental operations for which the physical process or outcome is secondary?

For example, performing a numerical calculation, understanding and selectively applying a rule or principle, and identifying objects by name. If the skill is essentially mental, the task should be considered for the written component.

4-4. CAN ENTIRE TASK BE TESTED FEASIBLY IN HOC? Does time, equipment, terrain, facilities, or availability of qualified scorers dictate the impossibility of testing the task in the hands-on component? If feasible to test the task in the hands-on component, it should be considered for this component.

4-5. CAN THE SKILLED PART BE TESTED FEASIBLY IN THE HOC? If the entire task cannot be tested feasibly in the hands-on component, perhaps you can evaluate one or more of the key skilled elements by changing the task limits.

For example, you cannot have soldiers fire the LAW in a hands-on component, but you can have them place the LAW in operation. All highly skilled physical elements that can be tested feasibly in the hands-on component should be considered for allocation to this component.

Other parts or elements of the task may be primarily mental or represent an unskilled requirement in the performance. These elements should be tested also and be considered for allocation to the written component.

If not feasible to test the skilled part of the task in the hands-on component, or to change task limits to allow meaningful testing, place the task in the PCC.

#### 4-6. COMPONENT LIMITS.

a. As tasks are allocated to components, some components may start to fill up. Remember the limits on the size of the components:

Hands-On: 0-16 SU for SQT 2 or 3; 0-10 SU  
for SQT 4 and 5.

Written: 30-50 SU.

Performance Certification: 0-10 SU.

TOTAL SQT: 30-76 SU

b. The purpose of the grouping (Chapter 2) was to insure that "most important" tasks were considered for testing first. When a component fills up during the allocation process, check each additional task against a task (usually the last one) already in the component and decide which is more important for testing in that component.

**Section II. PREPARE SQT PLAN**

**4-7.** After tasks have been allocated to components, the SQT Plan must be submitted to ITED. An outline for the content of the SQT Plan is attached (Figure 4-2).

**4-8.** If the Soldier's Manual for the MOS has not been printed, send a developmental draft of the SM with the SQT Plan.

A. Cover Sheet

1. Paragraph 1
  - a. MOS
  - b. SQT No.
2. Paragraph 2. Estimated number of people to be tested Army-wide in this skill level of the MOS.
3. Paragraph 3. Explain the system used to select tasks. Why and how were tasks selected for testing?
4. Paragraph 4. Explanation of why tasks were allocated to each component. Include justification for relief of responsibility for the hands-on component if one is not included.
5. Paragraph 5. Give name, grade, position of the person most directly responsible for development of SQT with AUTOVON number.
6. Paragraph 6. Give name, grade, position of the approving authority with AUTOVON number.

B. Inclosure 1, Test Plan

1. Paragraph 1 lists the number of scorable units by test component using ITED number system.
  - a. Written
  - b. Hands-on
  - c. Performance certification
2. Paragraph 2 lists the number of scorable units for each skill level being tested.

Figure 4-2. SQT Plan Outline

C. Inclosure 2, Performance Certification Component

List the task titles for the performance certification component.  
List the major areas and the units scored.

D. Inclosure 3, Hands-On Component

List the hands-on task titles grouped by major areas.

E. Inclosure 4, Written Component

List the scorable units in the written component test with the task title for each scorable unit. These also should be organized by major areas.

Figure 4-2. SQT Plan Outline, continued

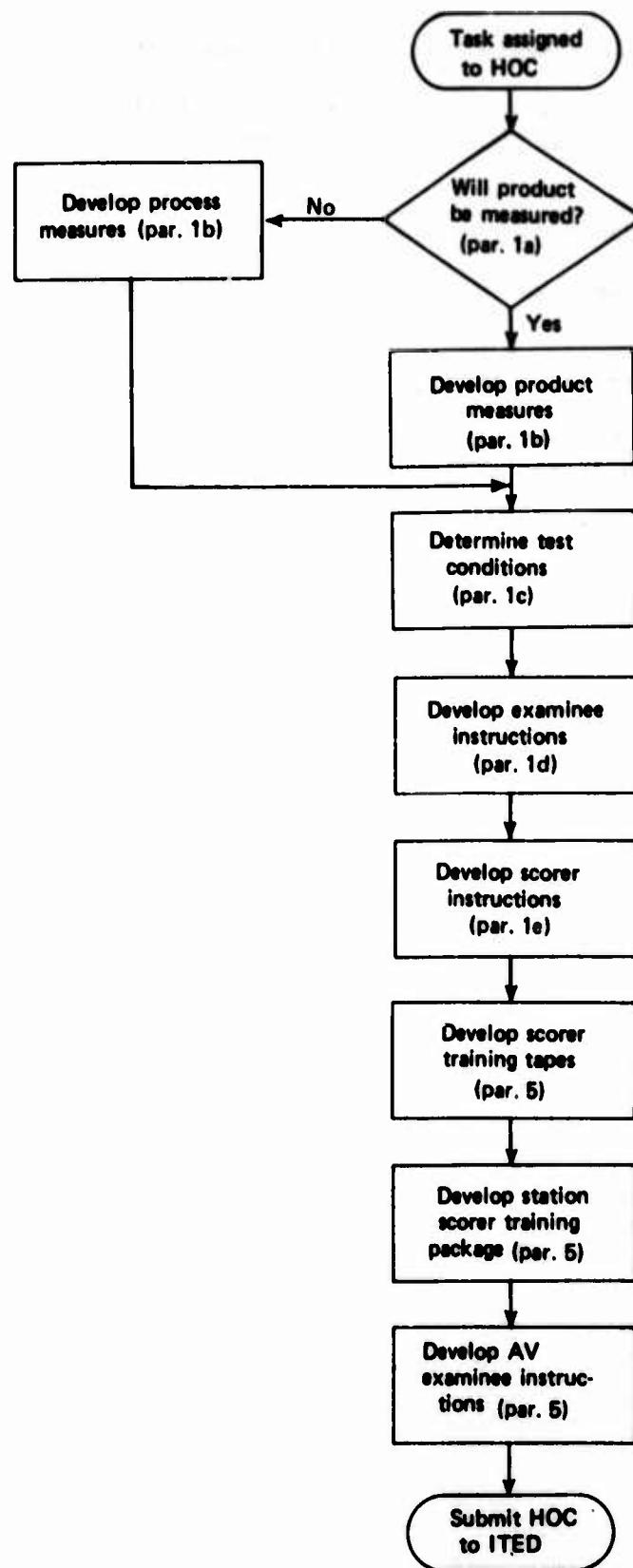


Figure 5-1. Hands-On Component Overview

## CHAPTER 5

## CONSTRUCT HANDS-ON COMPONENT SCORABLE UNIT

## OVERVIEW

The purpose of this chapter is to provide a procedure for constructing a complete hands-on SU. The end result is a standardized method of measuring a critical part of what the soldier does on the job.

1. Basically, five areas have to be developed in writing a hands-on SU. The five areas will answer five essential questions.

a. What is going to be scored?

This is answered by deciding whether the test will be process or product.

b. How will the soldier's performance be scored?

This is answered by developing process or product performance measures.

c. How will the test be set up?

This is answered by developing test conditions for the scorable unit.

d. How will the soldier know what to do on the test?

This is answered by preparing instructions to the examinee.

e. How will the scorer know what to do on the test?

This is answered by preparing instructions to the scorers.

2. As with any area of test development, the starting point is the task analysis. This will tell you:

a. What you are going to test (the task).

b. What the soldier must do to perform the task (elements).

c. How well the soldier must do the task (standards).

d. What will affect the soldier's performance (conditions).

3. Refine the information from task analysis into a hands-on test. During this procedure, you may find faults with the task analysis or information from the task analysis that will not fit your needs for hands-on development. These must be resolved as they arise.

4. Concentrate on the production of good reliable performance measures (product or process). During development and refinement of these performance measures, keep in mind what condition the equipment must be in, what guidelines the scorer will need, and what instructions the examinee will require.

If you do this and maintain an informal note system as you refine the performance measures, the test conditions, examinee instructions and scorer instructions will essentially be written by the time performance measures are completed.

5. As part of the hands-on package, each TDA will prepare scorer training tapes, a scorer training package for each test station and audio-visual tapes for examinee instructions.

6. As you go through this process, remember what you are trying to develop. That is:

- a. A test that measures what the soldier has to do on the job.
- b. A test that can be administered to a soldier in many different locales under standardized conditions.
- c. A test that can be scored fairly and accurately wherever and by whomever it is administered.

#### Section I. DETERMINE PRODUCT OR PROCESS

##### 5-1. INTRODUCTION.

a. The first step in developing a hands-on test is to determine whether it is a process- or a product-scored task. To determine this, decide what the scorer must observe to evaluate the soldier's performance. Must the scorer observe every element in the task performance or can the scorer focus on some product produced during the performance of the task? A simple example should serve to highlight the difference.

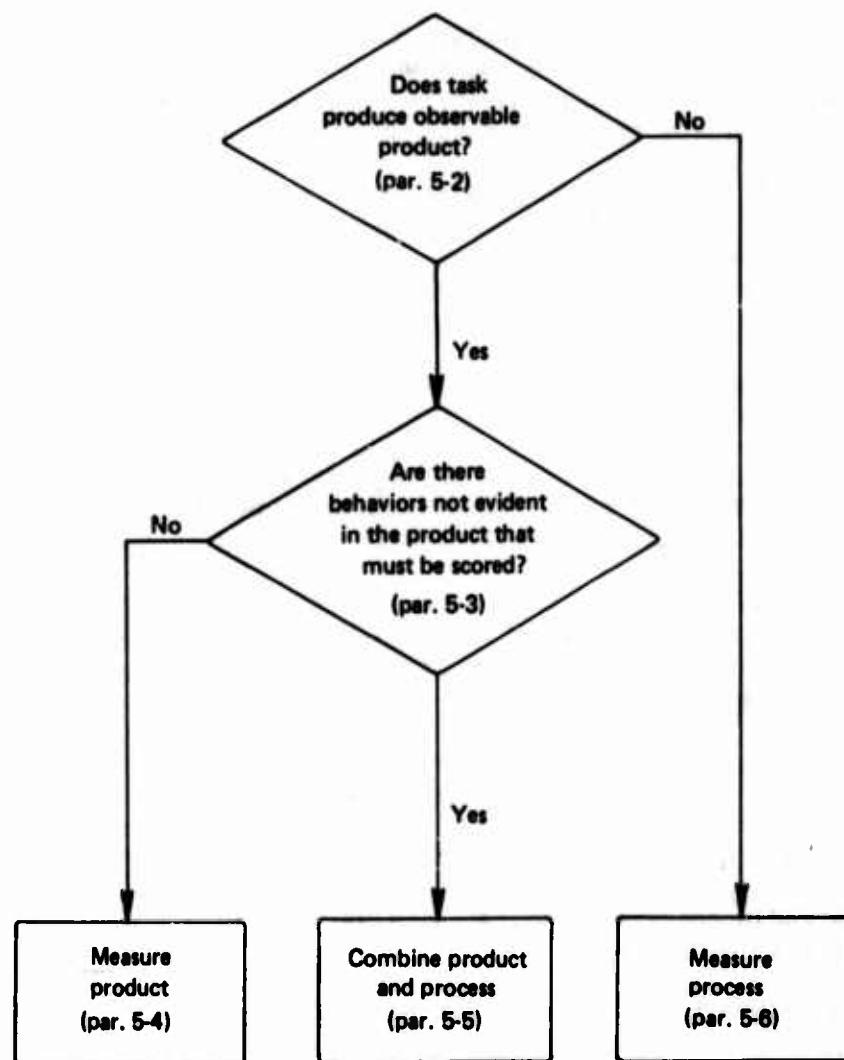


Figure 5-2. Determine Product or Process

Perform an about-face is a process task. (The scorer must observe the entire movement while it is being performed to determine adequacy of performance.)

Tie a square-knot is a product task. (The scorer would not even have to be present while the task is performed. Assuming adequate controls, the scorer can evaluate the performance anytime it is completed.)

b. The goal is to achieve product-type tests or measurements if possible, since they simplify test administration. Product scoring is generally more standardized, does not require detailed concentration on the part of the scorer and, for some tasks, can allow more than one individual to be tested at the same time.

c. Differences between product and process become more subtle as individual tasks are considered for testing.

Indeed you must understand the process of any task before a process/product decision can be made. Most tasks will be thought of initially as process tasks because they will appear that way in the SM.

A task is usually taught as a series of steps (elements); therefore developers may tend to think of most tasks as process. Now the emphasis is on testing, and many tasks that are outwardly process, in fact, can be product-scored during testing.

d. There is usually no problem in identifying products that are fabricated; that is, something produced something that we can hold in our hands. Most "products" of tasks, however, are not that distinctive or tangible. Many products will be actions or results. Some examples are:

A radio check (audio) for placing a radio set into operation.

A weapon function check for assembling a pistol.

**5-2. DOES THE TASK PRODUCE AN OBSERVABLE PRODUCT?** You must decide if the end result of a task meets the definition of a product. You should be able to state in a few words what the end result is and then decide if this is measurable as a product.

Products allow us to make certain assumptions about the process of their production even though this was not observed. If the task has no product, it must be scored as a pure process test.

5-3. ARE THERE ELEMENTS NOT EVIDENT IN THE PRODUCT THAT MUST BE SCORED? Even if you decide that the task has an observable product, examine the task analysis to identify which elements of the task are measurable in the product, and which are not. The following factors should be considered in identifying elements that cannot be scored from the product:

a. Safety factors. Safety will be as important a concern in testing situations as it is on the job. Safety factors include actions that could lead to personal injury or damaged equipment, and are identified as key elements in the task analysis.

Safety factors must be specifically stated. They require vigilance on the part of the scorer to prevent violations. Not all safety considerations will be included; only those identified as key elements.

Safety considerations become process-tested measures since they usually require the scorer to intervene in the test to prevent, if possible, violations from occurring.

b. Elements hidden by the product itself. Some process elements become physically obscured by the product or end result to the extent that they cannot be determined readily by outward examination of the product. Some product assembly or fabrication tasks are of this type.

More common are tasks that involve a product but which contain elements that cannot be evaluated accurately from the product because they are not lasting. For example, clearing the protective mask or turning off radios before starting a vehicle.

c. Diagnostic evaluation. You must anticipate the requirement for feedback of test results. "Pure" product tests will identify soldiers who cannot perform the task, but may not reveal which elements they fail most frequently.

Therefore, when considering a product test, you must decide if a requirement exists for diagnostic feedback. If there is, you must provide for scoring of elements in order to provide that feedback.

Often this will be a matter of policy and should be coordinated with Directorate of Evaluation or other group responsible for evaluating and reporting post-test information in the TDA.

5-4. MEASURE PRODUCT. If all elements can be measured accurately from the product or end result, the test will become a "pure" product test, and you need only describe the product. How to do this will be detailed later.

5-5. COMBINE PRODUCT AND PROCESS. Many tasks will not fall neatly into the product or the process classification. Therefore the term, "Combination Process/Product" will apply to many tasks.

In the combination of the two types, you first identify those elements that must be evaluated during the process of performance, then use product measures for the remainder of the task. Even in process tests, the scorer should be directed to elements where measurement is best made.

## Section II. SPECIFY PERFORMANCE MEASURES: PROCESS

5-6. INTRODUCTION. In this part, you must tell the scorer what you want him to look at and how you want him to judge what he sees. The starting point for process measures are elements as listed in the task analysis.

Now, refine those statements so they are presented in a way that can be scored accurately and consistently by many different scorers. Remember, in process-scoring we are looking at what the soldier is doing as he is doing it.

5-7. ARE PERFORMANCE STANDARDS STATED? For many elements (performance measures), it is insufficient to simply state the action because you believe the standard is implied in the element. During review of task analysis, you attempted to determine if the element was described in sufficient detail so that incorrect performance could be determined.

At this point, you are double-checking to see that the standard is specified for each performance measure. Standards answer questions relating to "how much" or "how well." You should look carefully and ask questions about terms like "turn," "adjust," "tighten," "loosen," and other words that imply judgment about whether the step is performed correctly.

Likewise, be wary of the "ly" adverbs in describing standards. Standards using terms such as "properly," "rapidly," "firmly," "accurately," are often unmeasurable and may be scored differently by different scorers.

## 5-8. IS MORE THAN ONE PASS/FAIL PERFORMANCE MEASURE STATED?

a. Each performance measure should spell out a distinct action. Avoid confusing the scorer. For example: "Removes clip and pulls slide to the rear," as a performance measure, may be confusing to a scorer. What about the soldier who removes the clip but does not pull the slide to the rear?

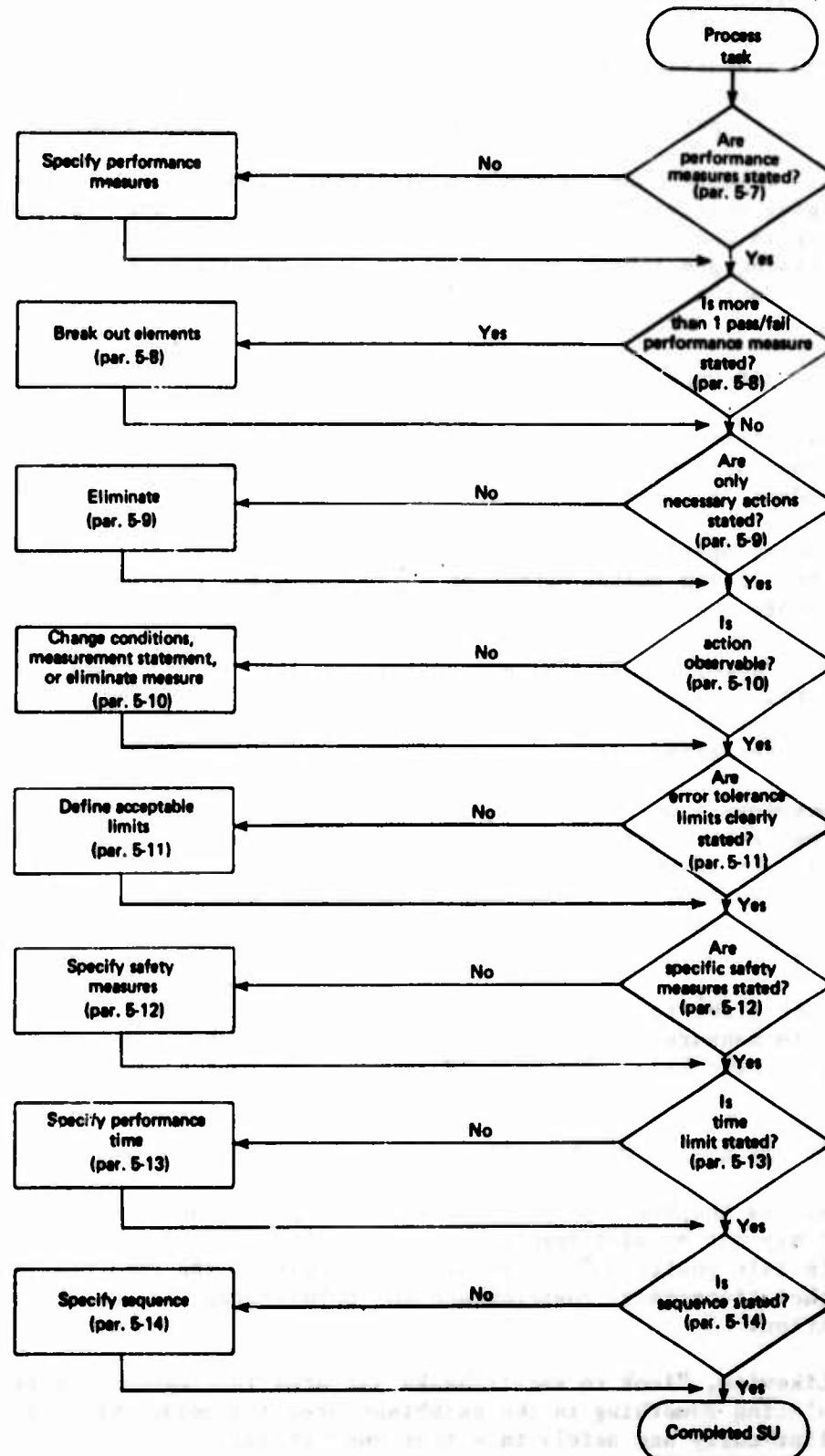


Figure 5-3. Specify Performance Measures: Process

You are asking the scorer to make a judgment on which was more important. Will he score it the way you want him to? To insure standardized scoring, separate any performance measures that are lumped together.

b. Simultaneous actions are different. "Depresses accelerator while engaging starter," is an acceptable statement. So are statements of how a step is performed, such as "alines sight by traversing, elevating or depressing gun tube." Often such statements will be needed to cue the scorer to the actions performed.

#### 5-9. ARE ONLY NECESSARY ACTIONS LISTED?

a. After breaking out statements into separate elements, you may have too many actions listed. Some actions need not be stated if they are dependent on other actions.

For example: "releases safety cover; activates firing switch." Assuming the firing switch cannot be activated without releasing the cover, eliminate the first element from scoring.

b. It may not be always necessary to list every element from the task analysis as a performance measure on the test. Remember the machine scoresheet is limited to pass/fail for 20 items. No exact rules can be stated for elimination of elements if over 20 are listed.

Developers must work with subject matter experts to direct the scorer to that point where measurement is best made but where no question remains as to whether a step was performed correctly or incorrectly. You must be careful not to overwhelm the scorer with performance measures; the scorer cannot observe the soldier while he reads and marks the scoresheet.

Some elements listed may be performed in the task but are not essential to measure. Try to direct the scorer to that point where measurement is best made. The goal is to make the test easy to score.

#### 5-10. IS ACTION OBSERVABLE?

a. Is the behavior observable as a physical act? An eye or head movement may not be discernable as a physical act. For example: "Throttle is in idle position," can be evaluated only if the throttle is placed in another (incorrect) position and the soldier has to move it to the idle position.

Likewise, "look to see if backblast area is clear" could be evaluated by placing something in the backblast area (assuming this can be done realistically and safely in a test condition).

Some caution is necessary when preparing the scoresheet. Sometimes you must use parenthetical notes to the scorer as to how he evaluates the measure of "checking."

b. Some ambiguous actions may be the result of poor wording of a performance measure. For example: "Checks radar scope" may actually mean the soldier has to turn and look at the scope. Cue the scorer to the performance by stating, "turns and looks at scope #1 before activating scanner."

5-11. ARE ERROR TOLERANCE LIMITS ALLOWED? Error tolerance limits allow a range of accuracy on some standards. Such limits are often needed in dealing with standards involving measurements. Avoid a specific number unless that number is the actual standard.

Most actions will allow a range; for example, "700-800 rpm," "within 50 meters," "+ 50." Specify the range allowed. If appropriate, specify how the measurement is made or what the mean reference point is.

5-12. ARE SPECIFIC SAFETY MEASURES STATED?

a. If safety is a factor, write a safety performance measure. This performance measure cites a specific action must or must not be done. The measure should cause the scorer to intervene in the testing to prevent the violation if possible (whether it should cause a NO-GO score is up to the TDA).

b. Statements such as "commits no safety violations," are unspecific, unstandardized and unacceptable.

5-13. IS TIME LIMIT STATED?

a. If the task is time-constrained, the time limit is a performance measure. For example, "Mask in 9 seconds." Insure the start and stop points are clearly identified, both to the scorer and the examinee.

Time-constrained tasks are identified usually as such in the task analysis. If you have altered the test limits, insure the time is adjusted accordingly.

b. If the task is not time-constrained, a time limit must still be assigned to enable efficient test administration. Utilize job experts and field tryouts to arrive at reasonable time limits.

c. As a final consideration, make an estimate of the time the scorables unit will take. Look at performance time, plus setup time, scoring time, and examinee briefing time. This is the total administration time.

You may have guidance on maximum allowable time for an SU. If not, look closely at any item exceeding 30 minutes. The entire component must be administered in four hours. If time estimates are accurate and the SU is too long, you may need to tighten the test limits some.

**5-14. IS SEQUENCE STATED?**

a. Examine the task elements carefully for sequence requirements. Some tasks require total sequence, others only partial sequence. If sequence is required, state it as a performance measure.

b. If no sequence is required, state this for the scorer on the scoresheet.

**Section III. SPECIFY PERFORMANCE MEASURES: PRODUCT**

**5-15. INTRODUCTION.** In this block, you must decide what the scorer needs to look at when he has a product test (either as a pure product test or included in a combination process/product test) in order to evaluate a soldier's performance.

Much of what was already discussed in process testing also applies here. The scorer is concerned here in looking at something a soldier has done, not in watching him do it. This should both ease the job of the scorer and make his evaluation more accurate.

But it places a greater burden to insure that you precisely define the correct outcome. Though the starting point is again task analysis, products often are not clearly evident or defined in the task analysis. Look for end results - not how they were achieved.

**5-16. DEFINE ACCEPTABLE PRODUCT.**

a. This equates with the standard of performance. It may have identifiable characteristics or it may be describable only by the absence of unacceptable characteristics.

For example, the standard (or acceptable product) for typing correspondence may be that it follows a specified format and contains no strikeovers or misspelled words.

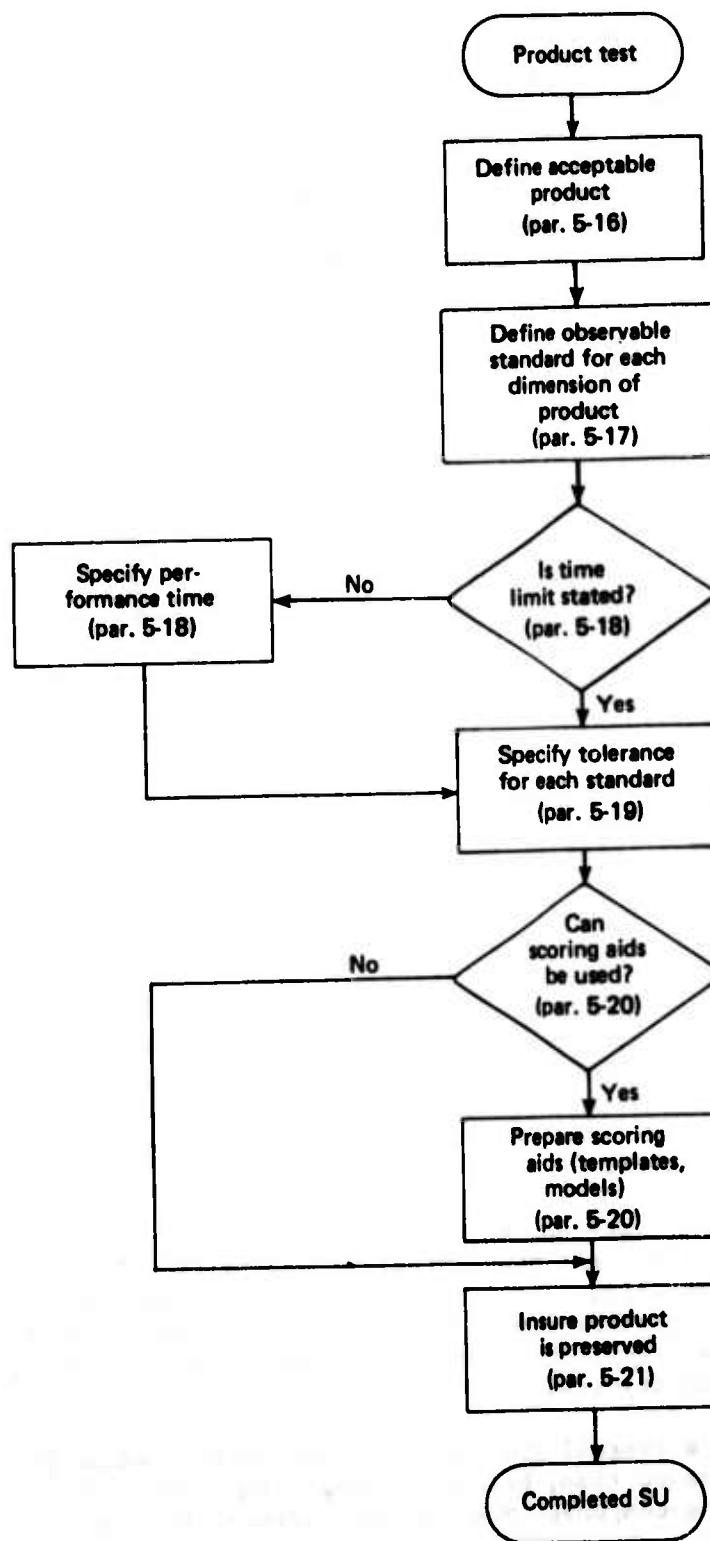


Figure 5-4. Specify Performance Measures: Product

b. Check for the standard in the task analysis, but remember it may not be readily apparent. Most task analyses are prepared in a procedural format. If you cannot deduce a standard for the product from the task analysis, consult with subject matter experts.

Remember an "acceptable product" may be a condition such as a running engine or ringing telephone; it may be an action such as a weapon function check; or something produced or constructed such as a completed form or a computer program measurable by description or by comparison with a model.

5-17. DEFINE OBSERVABLE STANDARD FOR EACH DIMENSION OF THE PRODUCT.

a. Carefully examine products for all dimensions and define all the standards necessary to assist in scoring. Examples of various dimensions are:

-height	-tightness	-shape	-accuracy
-width	-strength	-color	-time
-depth	-elasticity	-location	-speed
-length	-precision	-position	-size
-format	-tolerance	-sound	-adjustment
-texture		-rhythm	

b. Insure that each dimension is specified in a manner that is observable or measurable. Remember, most measures or standards must be suitable for rating by unsophisticated methods.

5-18. IS PERFORMANCE TIMED?

a. Some product tests are time-constrained. Time is usually specified as a standard in the task analysis. Time-constrained tasks usually will include the time as part of the product standard. Start-and-stop points must be clearly defined. The stop point (or limit) also must be clearly defined to the soldier.

b. If performance is not time-constrained, a time limit must be established to insure efficient test administration. This time limit is not meant to be discriminatory, i.e., the soldier who can perform the task does so easily within the time limit, the soldier who cannot likely fails for reasons other than time. Determine performance times with subject matter expert and field tryouts.

c. Finally, consider the overall time required for administering the scorable unit. Add performance time, briefing time, setup time, and scoring time. Remember, the component must be administered in four hours.

You may have received guidance on how long an SU can be. If not, look carefully at any SU where the total time exceeds 30 minutes. You may want to redefine the test limits to shorten the required time.

5-19. SPECIFY TOLERANCE FOR EACH STANDARD. If a range of measurement is acceptable in the standard, insure that the range is specified. Few products or end results that involve a numerical measurement involve a specified number. If a range is allowed, insure that range is specified; for example, "50-70 FPs torque," "within 50 meters."

5-20. CAN SCORING AIDS BE USED?

a. Scoring aids are used to help standardize what is being scored. Consider the use of the following kinds of aids:

- |             |                                   |
|-------------|-----------------------------------|
| -overlays   | -diagrams                         |
| -templates  | -photographs of correct product   |
| -blueprints | -photographs showing unacceptable |
| -models     | characteristics                   |

b. Scoring aids are not easy to prepare but should be considered where mere words do not convey an accurate understanding of the standard or where use of aids will speed up the scoring process. For example: the tasks "weld a bead," "prepare route overlay," "wire a circuit," are products difficult to describe in words but clearly shown by example.

c. If scoring aids are used, they must be produced in a quantity sufficient to test the MOS worldwide. Therefore, you must be sure that the aids are reproducible and can be readily distributed. Coordinate with ITED when scoring aids are planned.

5-21. INSURE PRODUCT IS PRESERVED.

a. The product need be preserved only until evaluated and a record (score) made. In most cases, this action will be immediately after production. However, some products do not last long, e.g., the ringing of a telephone or a communications check.

Others are easily destroyed, e.g., a sight alignment, a compression reading or a compass reading. Insure that instructions to the scorer include procedures in how and when to record scores for products of short duration or that are easily destroyed.

b. Some products may be preserved for later evaluation by tagging, recording, or even videotaping. Be sure a positive method for identifying each product is provided.

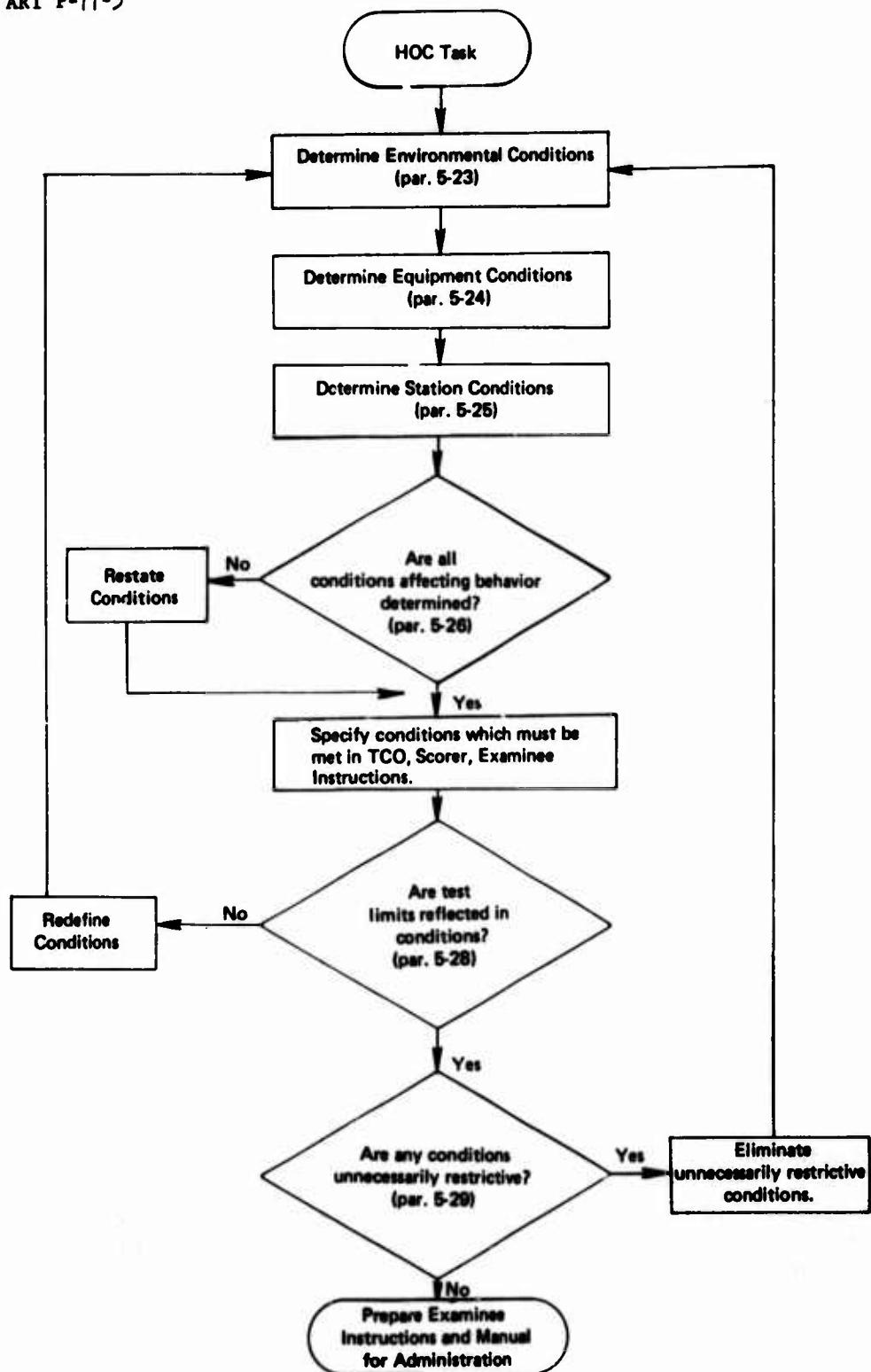


Figure 5-5. Determine Test Conditions

## Section IV. DETERMINE TEST CONDITIONS

## 5-22. INTRODUCTION.

a. In this block, you must determine what the soldier is going to experience (see, hear, feel) when he takes the test. We are concerned primarily with the effect these factors will have on his performance on the test.

In effect, you are telling the scorer how to set up the test so the soldier will take the test the way you visualized it. Start with the statement of conditions in the task analysis. However, task conditions will generally be too broad to meet the requirements for test conditions.

You must refine and redefine conditions until they insure standardized tests and eliminate variable influences on individual performance in the test situation.

b. The test conditions will be contained in detail in the Test Control Officers Manual for Administration. When the entire HOC is completed, conditions for an individual scorable unit may be consolidated into a station with other scorable units. Nonetheless, it is necessary to specify conditions for each scorable unit as developed.

Most of the information required for conditions should already exist in the notes made when you developed process/product measures. This step then requires only assembling those notes in a consolidated, logical format.

## 5-23. DETERMINE ENVIRONMENT CONDITIONS.

a. Environment conditions affect outdoor tasks much more than indoor tasks or tasks performed under cover. Normally, adverse environmental conditions will be the primary concern. As a minimum, consider the following environmental conditions:

-Light	-Visibility	-Temperature
-Precipitation		-Noise

b. Does the environment affect behavior? Often the task analysis will state that the task is "performed under all conditions." However, if one soldier does the task in the rain and another soldier in clement weather, will the test be fair?

Remember, we are talking about test conditions, not job conditions. When stating environment requirements, a general statement to avoid adverse conditions usually suffices, for example, during daylight hours with clement weather above zero Celsius.

c. Naturally, there will be geographical locations where environmental conditions cannot be met. Geographical extremes (Alaska in the winter; Canal Zone in the summer) probably will have to be considered as special conditions for test administration. Separate instructions will be addressed only to the TCO's in the geographical extremes involved. At this point, be concerned only with general instructions affecting the majority of locations in which this test will be administered.

**5-24. DETERMINE EQUIPMENT CONDITIONS.**

a. This is one of the most important aspects of hands-on testing. Consider what the soldier must be given in equipment when he arrives at the station for testing. It will not suffice to have just the equipment "there;" it must be in a specific position.

b. Consider first what equipment is needed to perform the task. List all items of equipment normally used on the job when the task is performed. Include manuals, JPA, or other references if they are used on the job.

Include all tools and ancillary equipment and materials such as cleaning material and POL products. Consider equipment needed by the scorer to score the test and to keep the station operating over a period of time.

c. Consider all aspects of the equipment itself. Specify settings and position of dials, valves, controls and switches. Tell the scorer, if necessary, how to build-in faults, and also how to eliminate unwanted cues.

d. Specify how to setup equipment for the first examinee. Also specify what to do in preparing equipment for the second examinee so it is laid out exactly as for the first examinee. Specify layout of materials and equipment used in conjunction with the major item.

**5-25. DETERMINE STATION CONDITIONS.**

a. Station conditions pertain to the area where the test is conducted. Station conditions may be incidental to the task, like space to work and an area conducive to testing.

Conditions, however, may be an integral part of the test, i.e., the task can be performed only if the station conditions are as specified. For example, boresighting a tank gun requires a line of sight of 1200 meters.

b. Station requirements depend on the task but consider as a minimum the following requirements:

-Indoor, outdoor	-Distance requirements
-Shop, motor park, workbench	-Terrain, vegetation

5-26. ARE ALL CONDITIONS AFFECTING BEHAVIOR DETERMINED? By this point most requirements to standardize test conditions should have been met.

However, review performance measures to be sure that conditions provide for both administration of the test and performance of the task by the examinee. If a question exists as to the possible course of the examinee's behavior (or what he will do), restate the conditions.

5-27. SPECIFY CONDITIONS IN THE ~~TCO~~-SCORER-OR EXAMINEE INSTRUCTIONS.

a. Translate conditions into specific instructions to people who conduct the test.

b. Describe condition of site, environment and equipment for the TCO/TSM to implement. Incorporate those statements in the test-specific Manual for Administration.

c. Include specific instructions to the scorer for setting up site, laying out equipment, and operating station in the Manual for Administration.

5-28. ARE TEST LIMITS REFLECTED IN CONDITIONS?

a. All conditions establish how much will be tested. Review start point or initiating cues and end point and insure these are reflected in the conditions.

For example, to test a soldier's ability to adjust a wheel bearing, we may not be concerned with testing the soldier's ability to jack up the vehicle and remove the hub. Therefore, the equipment condition specifies that the wheel is raised and blocked and the hub removed.

b. If limits are not reflected in the conditions, you may need to redefine or restate conditions. Insure that the soldier does not have to do more or less than you want to test.

**5-29. ARE ANY CONDITIONS UNNECESSARILY RESTRICTIVE?**

- a. Differentiate between what is really needed to insure standardized performance and what is "nice to have."

For example, shop stalls or maintenance bays enhance maintenance tasks, but they are not required for all maintenance tasks. Although 68°F is a nice indoor working temperature, it is rarely the only acceptable temperature.

- b. Can the conditions be met at test sites in the field? TDA often have more available equipment and better terrain than do field units.

An example is specifying the use of stopwatches which are unobtainable in any quantity in the field. While stopwatches are required for time-constrained tasks, they should not be required for all tasks.

- c. Remember, the more conditions you impose that the scorer or TCO cannot meet, the more likely he is to ignore the ones that are necessary.

**Section V. PREPARE EXAMINEE INSTRUCTIONS**

**5-30. INTRODUCTION.**

- a. During this block, you will decide what information is going to be given to the soldier when he arrives at the station to perform the task. The soldier must quickly and concisely comprehend what he is to do and mentally prepare himself for doing the task.

Remember, the scorer in the field will not be allowed to answer questions or expand on your instructions. He can only re-read your instructions to the soldier who has a question or doesn't understand what he is supposed to do.

- b. Ability to retain the spoken word is limited in all of us. Add to this the pressure that will exist for many soldiers in a test situation. Be concise and direct. Be clear, but don't "speak down" to the soldier.

Finally, remember that you are probably writing words that will be spoken or recorded. After you have prepared your instructions, read them aloud to a few colleagues. Do they know what you want them to do? Rework the instructions until they are clear, but do not confuse length of instructions with clarity.

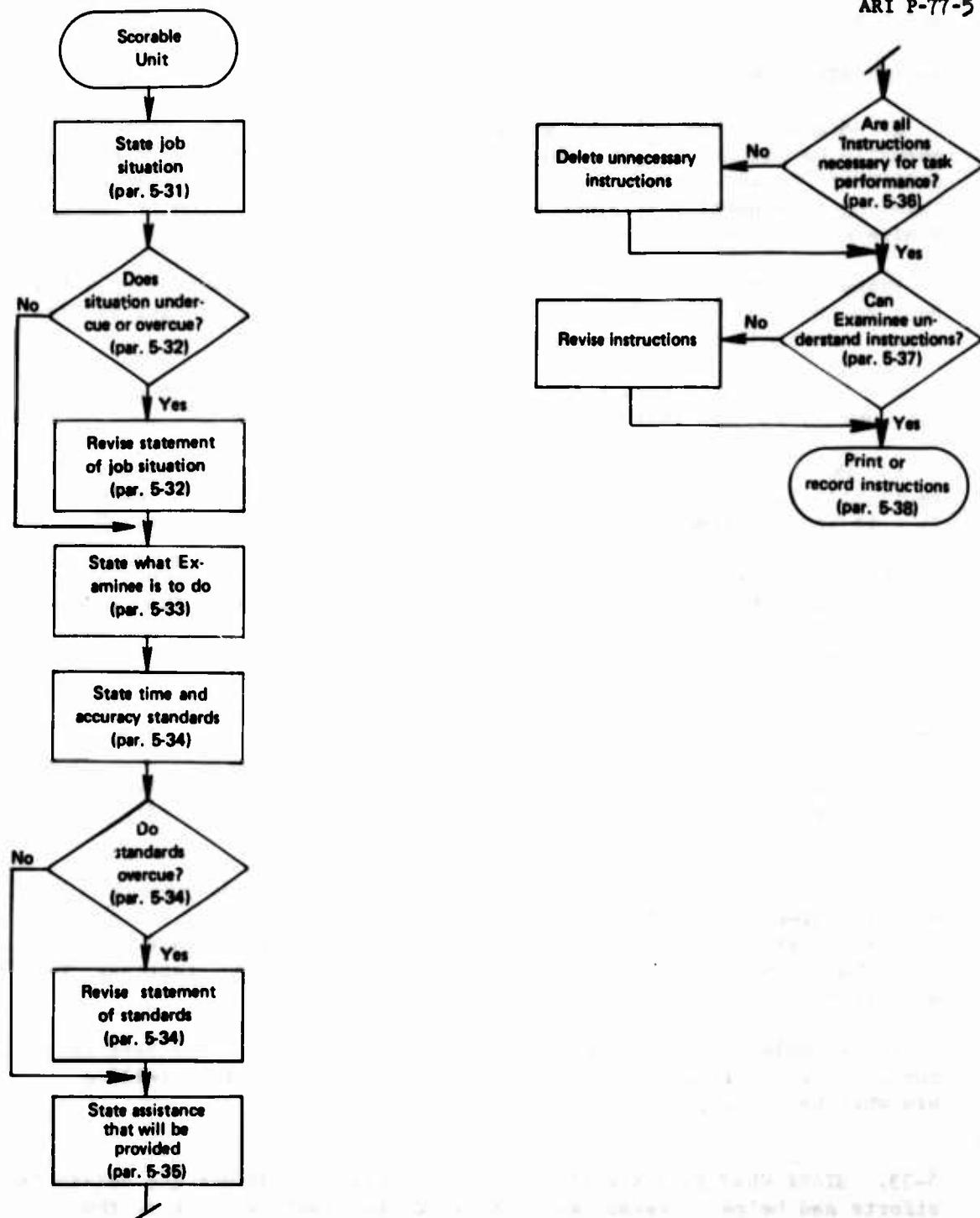


Figure 5-6. Prepare Examinee Instructions

5-31. STATE THE JOB SITUATION.

a. State the job role the soldier is in for this test. This will normally be his actual job. Some job situations require him to assume different "mental sets." It may be necessary to orient the soldier's thinking to a particular piece of equipment, a crew position, a function, a mission, or even a task.

Example: "You are the driver of an M113," "You are conducting a prisoner count at a confinement facility," "You are troubleshooting an aircraft electrical system."

b. Not all tasks need a job situation, e.g., masking; M16 assembly; but, unless you are sure, state the job situation.

c. Do not confuse job situation with unrealistic role playing. To test splinting a broken arm, tell him: "This soldier has a broken arm. It will be some time before you can get medical attention, etc..."

Do not involve the soldier with unnecessarily detailed situations: "You are a member of a long-range patrol returning from a night mission. A member of your patrol has fallen and broken his arm. You must move 5 km through enemy territory, etc..."

5-32. DOES SITUATION UNDERCUE OR OVERCUE?

a. Does it provide too little information (undercue) to start the task? Ask yourself, given this information what will be my first step? If not clear, give more information.

b. Does it tell the soldier what he is supposed to do when this is not our intent? Some tasks require that the soldier deduce his actions from the surroundings or conditions. Are you telling him more information than he would have at this point on the job? If so, rephrase the situation.

For example: "You are setting up the Claymore mine. You have not conducted the circuit check yet..." overcues the soldier by telling him what he is supposed to do.

5-33. STATE WHAT EXAMINEE IS TO DO. This helps to direct the soldier's efforts and helps to establish limits. State simply what it is the soldier is to do. For example: "You are to set up a Claymore mine for command detonation."

It may be necessary to include a statement on how far the soldier is supposed to go (limit) if the task is not being tested to its normal limits. For example: "It will not be necessary to replace the engine cover;" "Do not enter the control number in the log."

Orient the soldier toward the first step in the task in a way that approximates the situation on the job. Avoid the situation where the soldier could have several correct starts unless that is part of the task, e.g., troubleshooting.

b. Do not undercue. Generally avoid terms like "Take appropriate action" to begin the task unless it is clear from the situation what the soldier should do.

c. Do not overcue. Do not tell the soldier how he will perform the task. For example: "You are to put on the protective mask and clear it."

5-34. STATE TIME AND ACCURACY.

a. Always give the time limit.

b. Emphasize time-constrained task limits by including the time limit in the statement of what the soldier is to do. Example: "You must mask within nine seconds from when I announce 'GAS'."

c. Give the examinee the standard. Where a range of performance is acceptable or where it must be exact, tell him. Example: "Your final charge must be plus or minus 1/8."

d. Do standards overcue? Do not allow standards to overcue in the sense of revealing information on how to perform a step or task. For example: "After you return to the aiming point, the level bubble must be centered."

5-35. STATE ASSISTANCE THAT WILL BE PROVIDED.

a. If the task allows assistance, state exactly what that assistance will be. For example: "I will position the aiming stakes as you direct me;" "I will issue the fire command to you."

b. Except where obvious (for example in masking) state that there can be no assistance if none is provided.

5-36. ARE ALL INSTRUCTIONS NECESSARY FOR TASK PERFORMANCE? Examine the complete set of instructions. Is anything in the instruction not needed for task performance? Are terms explained that need no explanation? Are the instructions too long?

(Anything over four or five medium-length sentences, or over 15 seconds of total speaking time should be examined carefully.) Eliminate any portion of the instructions not essential to the task.

5-37. CAN THE EXAMINEE UNDERSTAND THE INSTRUCTIONS?

- a. This is a judgment. Is the terminology clear? Are the statements of what the soldier must do precise?
- b. Use terminology the soldier will understand. Do not say "prescribed load list" if the soldier calls it "PLL." Conversely, do not call it the M203 if the soldier calls it a "grenade launcher." Use both nomenclature and common names if there is doubt.
- c. Consider having the examinee repeat the instructions back to the scorer (not verbatim but enough to insure they are understood). At the very least, always ask the examinee, "Do you understand what you are required to do?" This gives the soldier a chance to hear the instructions again.

5-38. PRINT OR RECORD THE INSTRUCTIONS.

- a. Print the instructions to be read by the scorer or record them if you have an audio recording capability. If part of the instructions are given to the examinee for reference during the test, these must be extracted and printed on a separate reference card.

It is generally not effective to have the examinee try to read the instructions at the same time as they are being read to him, as the examinee then concentrates on neither.

- b. As part of the Manual for Administration, the TDA will prepare a brief set of instructions given to all examinees (probably in a group) before they start the test component. When preparing instructions for each scorable unit, consider what portions could be placed "up front" in the general instructions.

Try to present only instructions that apply to that specific scorable unit (or station). However, there may be some instructions that need to be repeated for emphasis at all or most of the stations.

Section VI. PREPARE SCORER INSTRUCTIONS

5-39. INTRODUCTION. To insure a fair, standardized hands-on test, it is necessary to develop specific instructions for the scorer for each scorable unit in the component.

In developing the test, you may discover some things that would assist in scoring the test. In short, you are probably now well qualified to perform as a scorer on this scorable unit.

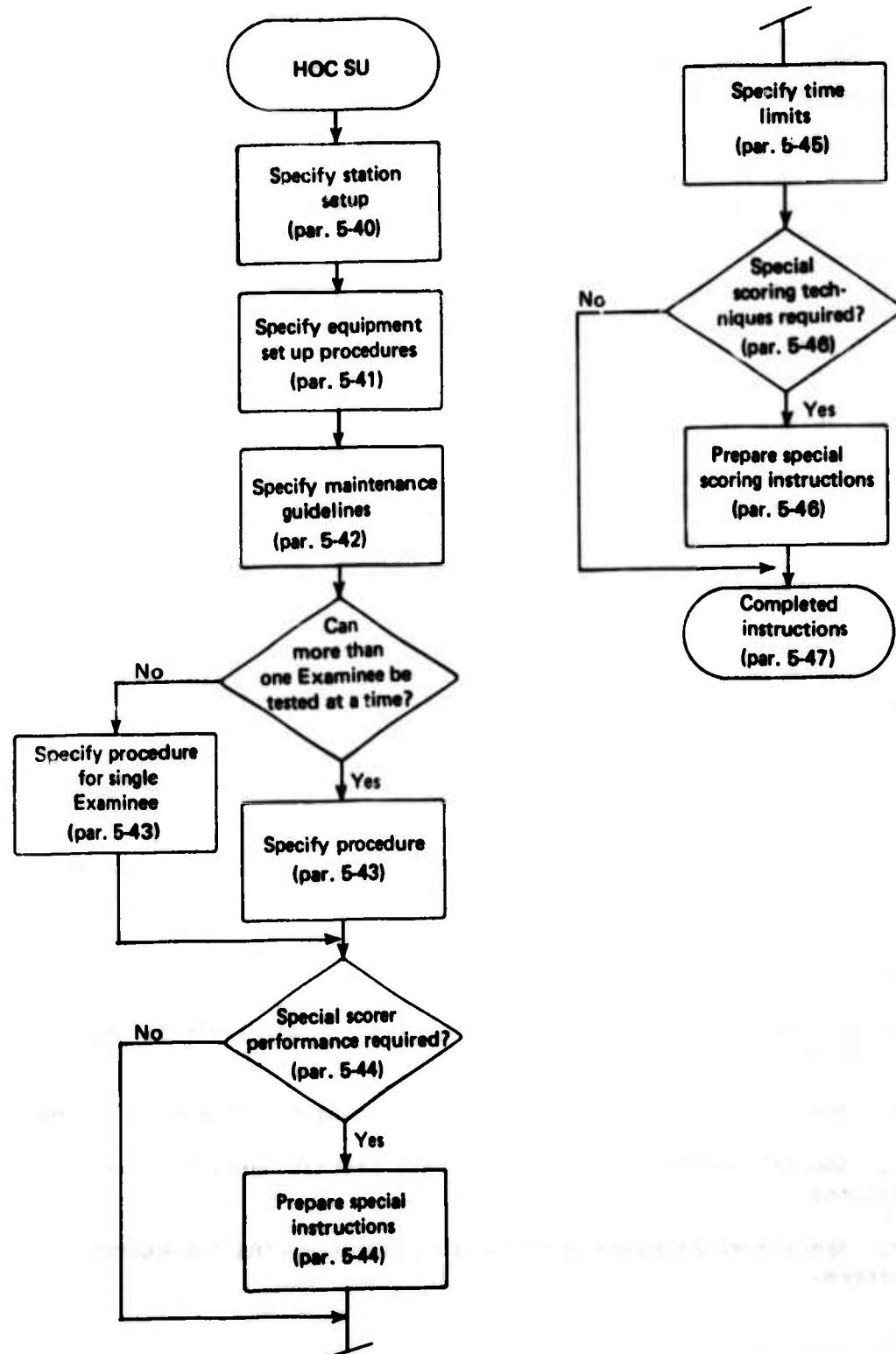


Figure 5-7. Prepare Scorer Instructions.

You must now impart this knowledge of how to score the test to an individual who has not had the benefit of your detailed analysis of the task. The scorer must be more than a soldier who marks PASS/FAIL on a scoresheet. He is a test station manager.

**5-40. SPECIFY STATION PREPARATION REQUIREMENTS.**

- a. The station requirements portion of the test conditions provides information to the TCO on selecting the area where the SU station will be located. It is necessary to detail specific instructions to the scorer on preparing the station for conduct of the test.
- b. For indoor tasks, you may want to insure the station is free from visual distractions that are present when the facility is in normal use. For outdoor tasks, you may instruct that all debris be removed, or the station be free of flammable materials.
- c. Station preparation requirements must not be beyond capabilities of one or two individuals to accomplish in a short period of time.

**5-41. SPECIFY EQUIPMENT SETUP PROCEDURES.**

- a. State in step-by-step procedures precise activities that scorer must perform before each examinee is tested. Include methods of controlling all variable conditions which affect standardization.
- b. Specify settings of dials, knobs or adjustments; position of equipment and tools, manuals and other material used in conjunction with the test.
- c. Use schematic drawings or layout diagrams to specify the setup of equipment.

**5-42. SPECIFY MAINTENANCE GUIDELINES.**

- a. Specify maintenance to perform at regular intervals during test administration.
- b. Specify maintenance to perform before each examinee is tested.
- c. Specify maintenance required under certain unpredictable conditions.
- d. Specify maintenance required to prevent cueing subsequent examinees.

5-43. CAN MORE THAN ONE EXAMINEE BE TESTED AT A TIME?

- a. When large numbers of soldiers are tested, it is desirable that more than one individual be tested at a time at a station. (NOTE: This is possible only when product tests are being used.)
- b. Examine the scorable unit to determine if this is possible. If it is, you must determine how many a single scorer can test feasibly at a time. (Normally, this will not exceed four.) In testing more than one examinee at once, you will have to decide how to position them; you may also have to change examinee instructions, or provide other special instructions to the scorer.
- c. Whether group testing is feasible or not, you must still specify where the examinee(s) should be at the start of the test, when and how he is to be provided information during the test, when he should be allowed to see the equipment layout, and how to handle examinee questions.

5-44. DOES TASK PERFORMANCE REQUIRE SPECIAL SCORER PERFORMANCE?

- a. If the scorer or assistant scorer is allowed to assist the examinee in any manner during the test, specify extent of the assistance.
- b. If, because of safety or other considerations, the scorer must interfere with the examinee's performance, tell the scorer when and how to intervene in the test.

5-45. SPECIFY TIME LIMITS. State specifically when the scorer must start to time the task and when he must stop. This is especially important in time-constrained tasks. The points must be easily identifiable.

5-46. ARE SPECIAL SCORING TECHNIQUES REQUIRED?

- a. Specify how a scorer must measure standards if this is not readily apparent. For example, the task of emplacing the M18A1 requires the scorer to look through the sighting window to determine alinement with the aiming point.
- b. Specify when and how to use scoring aids (templates, models, photos) if they are provided.
- c. Specify checkout procedures on equipment to eliminate false NO-GO caused by equipment failures.

5-47. PRINT SCORER INSTRUCTIONS. Scorer instructions pertaining to setting up equipment, maintenance and special situations are prepared in a special instruction sheet. Operation of the station must be specified in the Station Annex to the Manual for Administration.

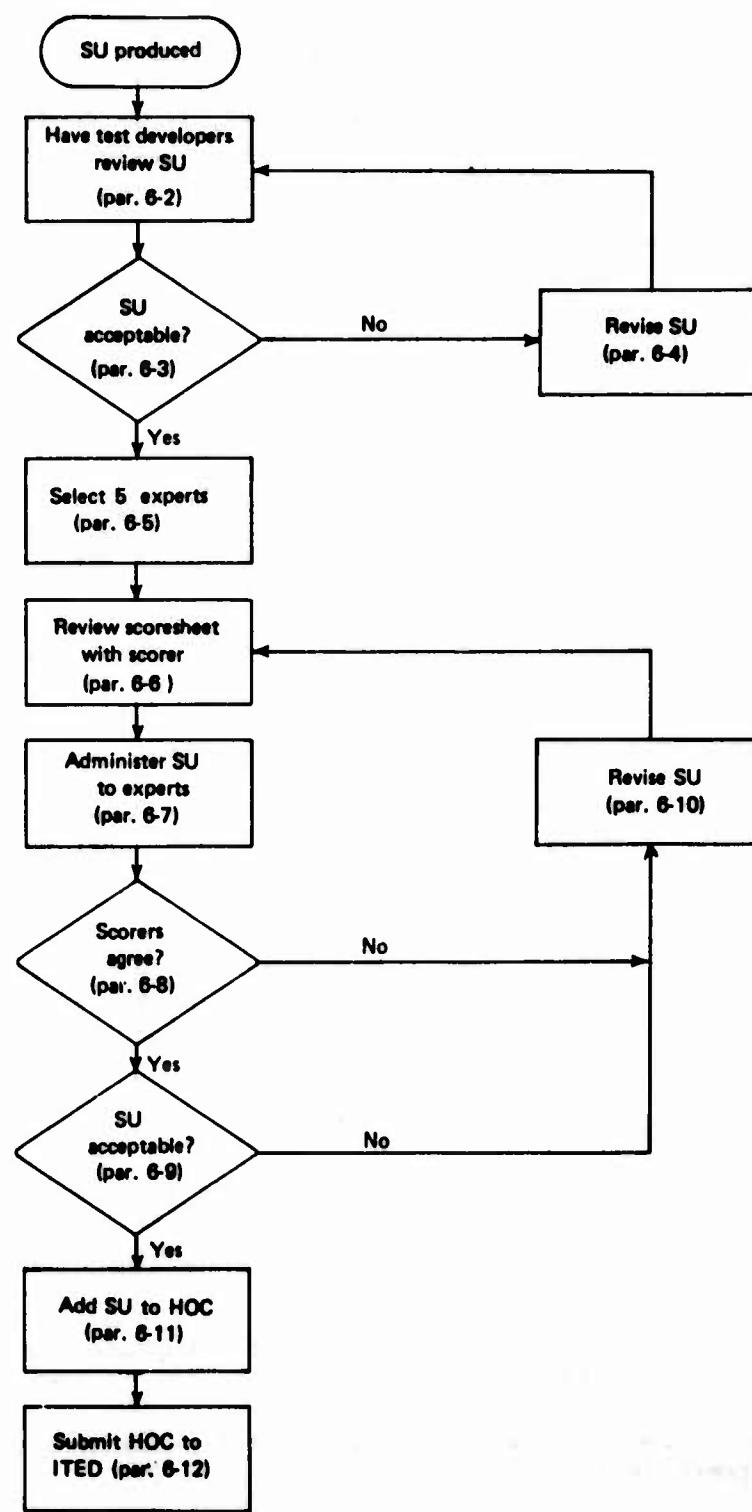


Figure 6-1. Try out Hands-On Component SU

## CHAPTER 6

### VALIDATE HANDS-ON COMPONENT

---

#### Section I. TRY OUT HANDS-ON COMPONENT SU

6-1. INTRODUCTION. This tryout is a systematic way to have experts in the MOS review the HOC to be sure each SU is consistent with doctrine. You must test experts and collect their opinions of the SU before you submit the hands-on component to ITED.

6-2. HAVE TEST DEVELOPERS REVIEW SU. Before you begin testing experts, have two or three people who have helped develop an SQT review the score-sheets and instructions for the SU. The test developers do not have to be qualified in the MOS covered by the SQT.

6-3. IS THE SU ACCEPTABLE? Ask each test developer to suggest changes in the SU. Have them look for performance measures that are not observable, check to be sure the standard and performance measures are described in enough detail for scorers to score the SU reliably, and check the clarity and completeness of the instructions.

6-4. REVISE. Weigh the suggestions of other test developers. If a suggestion makes sense, make the recommended changes. Have the test developers review the revised SU.

#### 6-5. SELECT AT LEAST FIVE EXPERTS.

a. In this case "experts" are people who you are sure can do the task covered by an SU. They do not have to be in the skill level for the SQT. In fact, experts will usually be senior NCO's or instructors in the MOS. An officer also may act as an expert for this tryout. One person may be an expert for several tasks covered by an SQT. Test developers should not act as experts for this part of the tryout.

b. The quality of this tryout depends on the level of expertise of the soldiers who score or take the SU. Do whatever you can to get high-level performers for the tryout. If you have any doubt about the qualifications of a potential expert, interview his peers and supervisor.

Ask the peers and supervisor how well the expected expert can perform the specific task. If they say the expected expert can do the task very well, consider including him in the tryout. If they say the expected expert can do the task only fairly well or worse, do not include him in the tryout.

c. After you locate several people you expect to be experts, give them a copy of the task analysis data for the SU and have them rate their ability to perform the task as described in the task analysis data. When you collect the self-ratings, emphasize the importance of accurate self-ratings.

Be sure expected experts know there is no benefit to them or to the SQT if they overrate their ability. Use one of these questions for the self-ratings. The question to choose depends on whether the task is performed normally under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well.
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

For this tryout, experts are people who choose No. 3, "Very well" or "No guidance."

6-6. REVIEW SCORESHEET WITH SCORERS. Designate two of the experts to act as scorers. Go over the scoresheet with the scorers to be sure they understand and agree with the performance measures and standard.

6-7. ADMINISTER SU TO EXPERTS. Test the remaining three experts one at a time on the SU. Lay out the equipment and set up the station the way it will be when you test for record. Have one scorer read the instructions. Have both scorers score each expert who is tested.

6- 8. DO SCORERS AGREE ON RATINGS? The first minimum requirement of an SQT hands-on SU is that each performance measure can be scored consistently. If the scorers disagree on any performance measure, discuss the disagreement with the scorers and the expert who was tested.

The disagreement may have resulted because one scorer was necessarily out of the recommended position for observing the procedure. If the disagreement resulted because of interpretation of the scoresheet, change the scoresheet or scorer instructions to eliminate the source of the disagreement.

6- 9. SU ACCEPTABLE?

a. The second minimum requirement of an SQT hands-on SU is that people who take the SU think it is a fair measure of their ability to do a critical part of their job. After an expert is tested, ask him or her the questions in Figure 6-2. Ask the scorers the same questions after everyone has been tested. If more than one of the experts or scorers think that the task is non-critical or that the test is unfair or incomplete, revise the SU.

b. Pay special attention to performance measures experts fail. Be sure that experts who fail a performance measure agree that their performance was incorrect. If they do not agree with the statement of the performance measure, check the accuracy of the task analysis data.

c. Also watch for "extra" things the experts do. Discuss those behaviors with the expert and scorers to be sure they are optional behaviors. If the experts claim they are essential behaviors, consider developing performance measures for them.

6-10. REVISE THE SU. Your primary concern during this tryout is that the SU is consistent with doctrine. If two scorers disagree about the correctness of an expert's performance, check the wording of the performance measure and completeness of the instructions to the scorers.

If experts disagree with the standard for any performance measure or for the SU, consider changing the standard. If experts think the task is not critical to the skill level addressed by the SQT, convey their reservations to the person who selected the task for testing. Have the same experts review any revisions.

6-11. ADD THE SU TO THE HOC. A hands-on SU that scorers score consistently and that experts consider acceptable meets first requirements for inclusion in the SQT.

6-12. SUBMIT HANDS-ON COMPONENT TO ITED. Submit the following items to ITED:

- The scoresheet for each SU.
- Administrative instructions.
- The task analysis data for each task.
- The SQT notice for the hands-on component.

Submit the HOC at the same time you submit the written component.

Once the HOC has been approved by  
ITED the content of each performance  
measure within each SU is fixed and  
cannot be changed without approval  
from ITED. Performance measures  
found to be unsatisfactory during  
validation will be revised to  
improve them, but not to the extent  
of changing content. Unsatisfactory  
performance measures cannot be  
deleted without approval.

NAME \_\_\_\_\_

**EXPERT AND RATER QUESTIONS FOR HANDS-ON SCORABLE UNITS**

Please answer the following questions for each Scorable Unit immediately after it is administered.

- |  | Yes | No |
|--|-----|----|
| 1. Is the task covered by the Scorable Unit a critical job requirement at the skill level? | —   | —  |
| If not, why not?   |     |    |
| 2. Does the Scorable Unit provide a fair measure of job requirements?                      | —   | —  |
| If not, what changes are required?   |     |    |
| a. Add Performance Measures?<br>Which ones?  |     |    |
| b. Delete Performance Measures?<br>Which ones?   |     |    |
| c. Modify Performance Measures?<br>How?  |     |    |
| d. Alter standards?<br>How?  |     |    |
| e. Clarify instructions?<br>How?   |     |    |
| f. Change test site or equipment requirements?<br>How?                                     |     |    |
| 3. Is the Scorable Unit administratively feasible?   | —   | —  |
| If not, what changes are required?   |     |    |
| 4. Do you have any additional comments or observations on this Scorable Unit?              | —   | —  |
| If yes, indicate them below.   |     |    |

**Figure 6-2. Question Sheet**

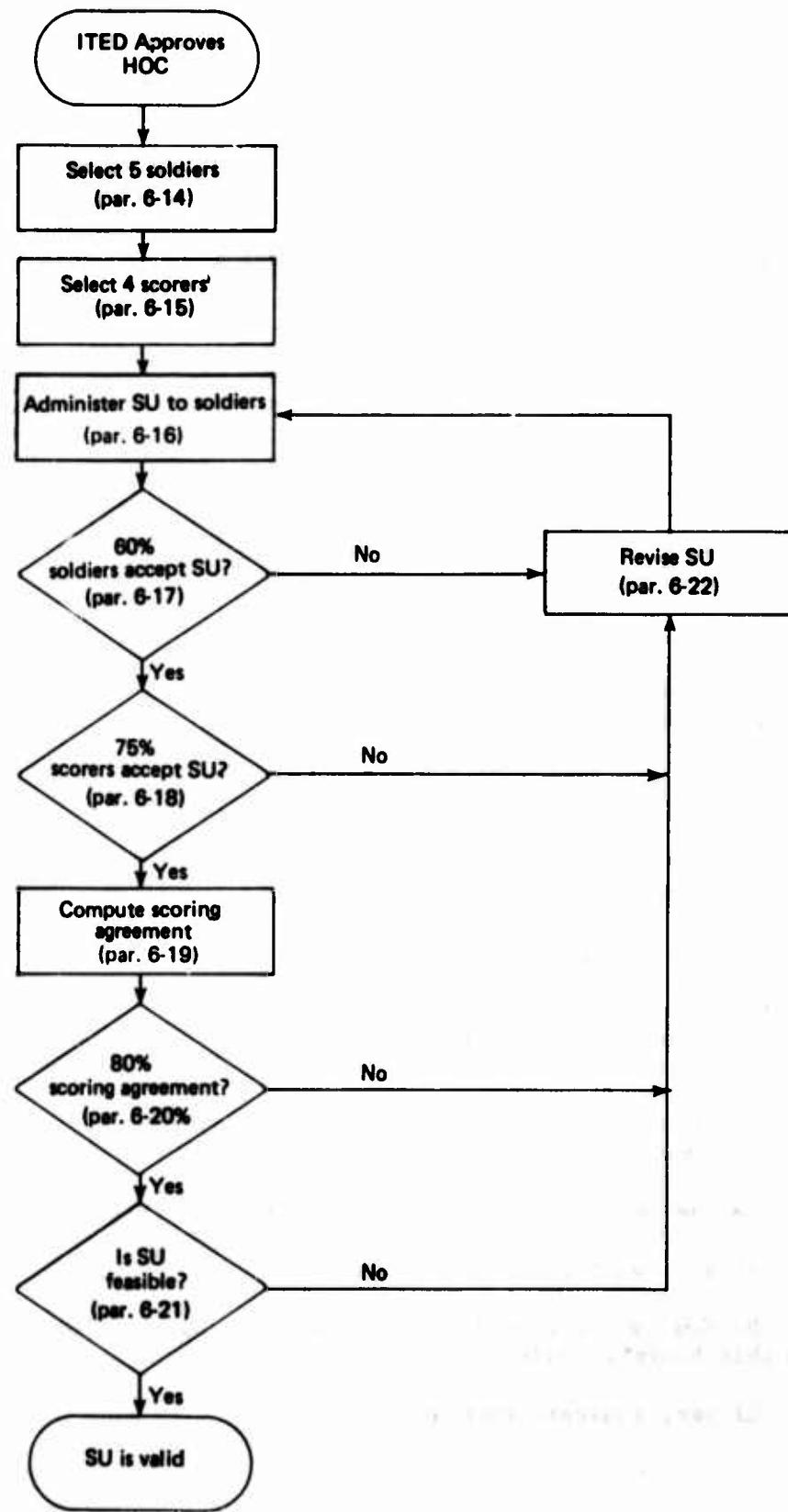


Figure 6-3. Validate Hands-On SU'

Section II. VALIDATE HANDS-ON SU

6-13. INTRODUCTION. This part of the validation will demonstrate whether scorers can score the hands-on SU reliably when representative MOS incumbents take the test. It will also give information on the acceptability of the SU to the representative incumbents.

6-14. SELECT FIVE SOLDIERS.

a. The validation sample should be composed of soldiers at the MOS skill level for the SQT. If several skill levels will take the SU, try to have each skill level represented in the validation sample.

b. Have all candidates for the sample rate their own ability to perform tasks covered by the written component. When collecting self-ratings, emphasize importance of accuracy. Be sure soldiers know there is no benefit to them or to the SQT if they overrate their ability. Give each soldier a detailed description of the task that specifies the task conditions and standard. Then, ask each soldier one of the following questions. The question to ask depends on whether the task is performed normally under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well.
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

c. If possible, select five soldiers who choose alternative 1 or 2. Have as few soldiers who choose alternative 3 as possible. Do not select anyone who chooses alternative 0, "Not familiar with the task."

6-15. SELECT FOUR SCORERS.

a. Scorers for the validation should meet the same requirements as scorers for the actual SQT. They should hold an equal or higher rank than soldiers in the sample and hold the MOS the SQT addresses. Ask potential scorers the questions in paragraph 2. Select four scorers who choose alternative 3.

b. Conduct a rehearsal for the scorers by having them act as examinees for each other. Watch each scorer administer the SU at least once.

6-16. ADMINISTER SU TO SOLDIERS. Test soldiers one at a time and have the four scorers score independently of each other. Set up the test conditions as much as possible like they will be when you test for record. Measure the time it takes each soldier to perform the task.

6-17. DO 60% OF SOLDIERS ACCEPT?

a. Ask soldiers these two questions as soon as they finish the SU:

(1) Do you think this task is an important part of your MOS?

(2) Do you think this test is a fair way to find out if you can do this task?

b. If any soldier answers "No" to either question, probe to find the reason for this answer. If the test is at fault, consider changing some characteristic of the test situation. Each question must be answered "Yes" by three soldiers to be considered acceptable. If more than two soldiers do not accept the SU, you must revise it.

6-18. DO 75% OF SCORERS ACCEPT?

a. After all soldiers have been tested, ask the scorers these two questions:

(1) Do you think this task is an important part of this MOS?

(2) Do you think this test is a fair way to find out if a soldier can do this task?

b. If any scorer answers "No" to either question, probe to find the reason for his answer. If the test is at fault, consider changing some characteristic of the test situation. If more than one scorer does not accept the SU, revise it.

## 6-19. COMPUTE SCORING AGREEMENT.

a. Collect scoresheets from the scorers. Then prepare a summary of scores for each performance measure. The summary should show the score each scorer gave each soldier on the performance measure. Here is an example of one form of such a summary:

## EXAMPLE OF A SATISFACTORY PERFORMANCE MEASURE

		Scorers			
		1	2	3	4
Soldiers	1	1	1	1	1
	2	1	1	1	1
	3	1	0	1	1
	4	1	1	1	1
	5	0	0	0	0

In the table a "1" means the scorer gave the soldier a "pass" on the performance measure; a "0" means the scorer gave the soldier a "fail" on the performance measure.

b. Next prepare a table that shows how many times each pair of scorers agreed. Here is a suggested format for such a table:

		Rater Pairs						Total Agree	Total Disagree
		1&2	1&3	1&4	2&3	2&4	3&4		
Soldiers	1	A	A	A	A	A	A	6	0
	2	A	A	A	A	A	A	6	0
	3	D	A	A	D	D	A	3	3
	4	A	D	A	D	A	D	3	3
	5	A	A	A	A	A	A	6	0
								24	6

An "A" means the scorers agreed; "D" means the scorers disagreed.

c. Finally divide the number of agreements by the number of possible agreements (agreements plus disagreements). In this case, there are 24 agreements and 30 possible agreements; therefore, the agreement is 80%.

d. Repeat the procedure for each performance measure in the SU.

6-20. EIGHTY PERCENT (80%) SCORING AGREEMENT. The minimum requirement for scoring agreement is 80 percent. This means that a scorer can disagree with the other three scorers no more than twice. A lower percentage indicates a potentially serious level of unreliability.

6-21. IS THE SU FEASIBLE?

a. Consider three questions in making this decision:

- Is the time limit reasonable?
- Will the length of this SU allow the installation with the most people in the MOS to conduct the component in five days?
- Are the test conditions essentially the same for each soldier?

b. If the answer to any of these questions is "No," revise the SU to increase its feasibility.

6-22. REVISE THE SU.

a. If the people who developed the SU have followed the development procedures, minor revisions probably will make the SU valid. That is good, because there is not much you can do to revise a hands-on test that has come this far in the validation procedure.

b. If soldiers or scorers do not accept the SU, consider changing the test conditions or changing the scope of the task. Soldiers may have resources on the job not available at the test site. If so, make those resources available.

Sometimes part of the task will be done by someone else. Eliminate that kind of irrelevant requirement. If you change the content, coordinate with ITED.

c. If scoring agreement is too low, the performance measures may not be stated clearly. Rework such performance measures. It is also possible that one of four scorers must stand where he cannot see all of the performance.

Consider that possibility before making major changes to the SU. If the disagreement cannot be attributed to a superficial weakness, contact ITED about the possibility of dropping the performance measures from the SU.

d. If you revise the SU, try it out again. You may use the same sample of soldiers and scorers if they are available. If the same soldiers and scorers are not available, select a new sample and repeat the data collection and analysis.

### Section III. CONDUCT FIELD TRYOUT OF HOC

6-23. SELECT APPROPRIATE UNIT. The last step in validating a hands-on component is to determine its administrative feasibility in a unit. FORSCOM, Army Reserve or National Guard units may be selected for the site of the field tryout. Coordinate through ITED with the major command to select a unit. The unit selected should have at least ten, and preferably up to 50 soldiers available for testing. The intent is to exercise the component in a realistic environment.

If you must test large numbers in a brief time operationally, have a large sample for this tryout. If you must test only a few soldiers operationally, you may have fewer than 10. The soldiers must be in the MOS skill level that will take the hands-on component.

6-24. SEND TEST MATERIALS TO TCO. Send the TCO who will conduct the test the following materials:

- Twice as many scoresheets as the number of soldiers to be tested. (The scoresheet will serve as an SQT Notice to the soldiers who will be tested.)
- A list of the SU to be tested at each station.
- Instructions for setting up the test site.
- Instructions for setting up and maintaining each station.
- Instructions for scoring each SU.
- Requirements to be a scorer of SU.

6-25. OBSERVE TESTING. Send a representative from the TDA to observe the testing.

ARI P-77-5

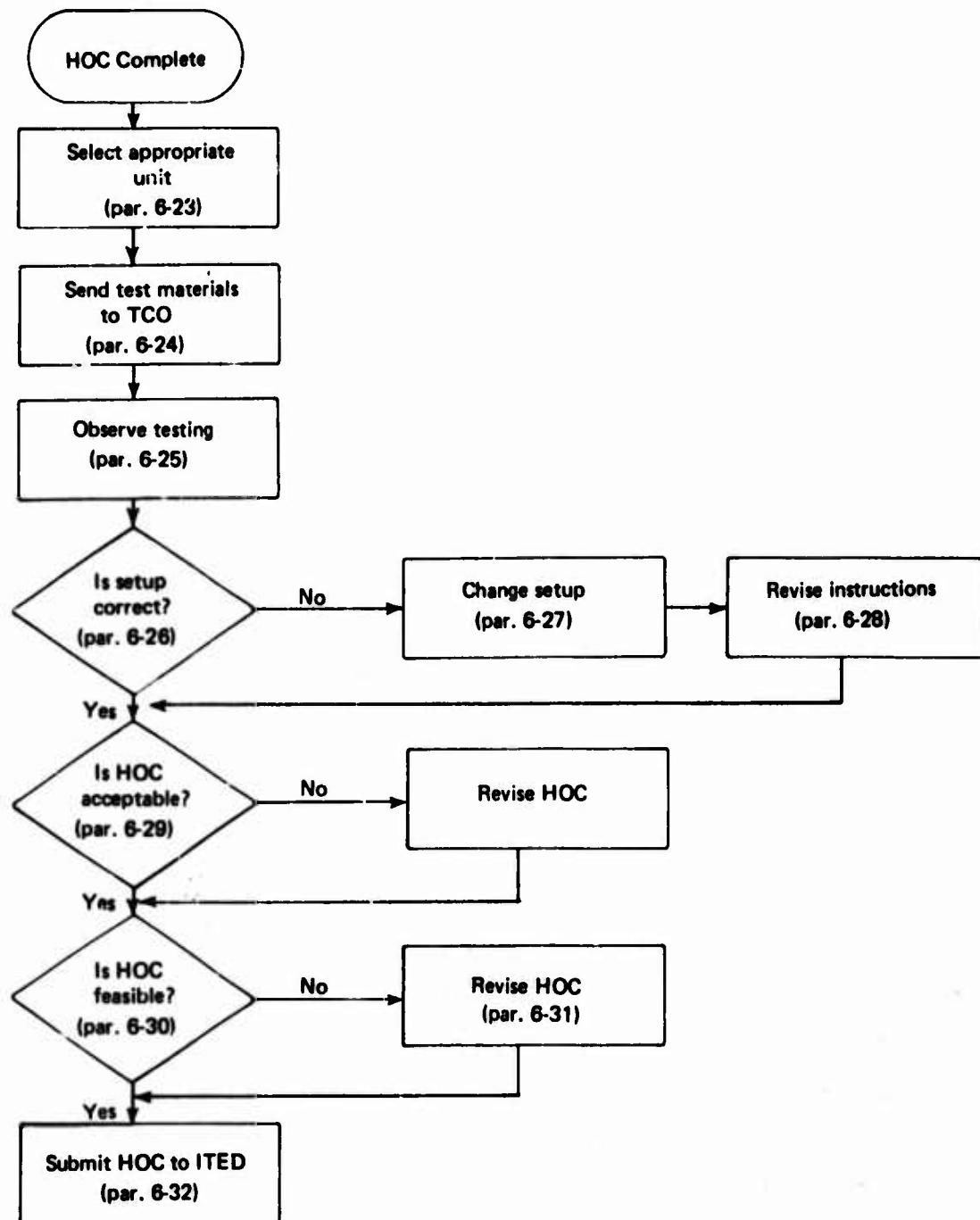


Figure 6-4. Conduct Field Tryout of HOC

6-26. IS SET-UP CORRECT? Be sure the TCO and the scorers have followed instructions for setting up test site and individual test stations.

6-27. CHANGE SET-UP. If the setup is different than the instructions call for, learn why. If the new setup is an improvement, leave it alone. If the new setup might make test conditions different than planned, have the TCO or scorer change the setup to conform to requirements.

6-28. REVISE INSTRUCTIONS. If an incorrect setup is caused by unclear instructions, clarify them. If a new setup is an improvement, change instructions so they call for the new setup.

6-29. IS HOC ACCEPTABLE? Have each soldier and scorer fill out Question Sheet 2 (Figure 6-2). Have each soldier who is tested fill out Question Sheet 4 (Figure 6-5). If more than 20 percent of the soldiers and scorers think any SU is unfair, review conditions, standard and performance measures of that SU for possible revision.

6-30. IS HOC FEASIBLE? This tryout is your only chance to find out for sure how many scorers and how much equipment will be required to test a given number of soldiers on the HOC. Pay special attention to amount of time each scorer takes to administer each SU at the station to one soldier.

Based on those figures, project the number of setups of each station you would need to test the largest number of soldiers in the MOS at one installation within five days. Decide whether equipment and scorer requirements are reasonable.

6-31. REVISE HOC. If requirements are not feasible, coordinate with ITED to reduce the scope of, or drop, the troublesome SU.

6-32. SUBMIT HOC TO ITED. Send revised scoresheets, revised instructions and summaries of the validation and field tryout to ITED.

NAME \_\_\_\_\_

QUESTION SHEET 4

Do you accept this test as a fair measure of your ability to perform  
in your MOS?

	Yes	No
Test 1	—	—
Test 2	—	—
Test 3	—	—
etc.	—	—

If not, please explain why.

Figure 6-5. Examinee Questions for Hands-On Scorable Units

## CHAPTER 7

## CONSTRUCT WRITTEN SCORABLE UNIT

## OVERVIEW

1. OBJECTIVE. The objective of the Written Component (WC) is to test the soldier's ability to perform a task by evaluating his responses to written questions. The objective should not be thought of as just a test of knowledge, since all questions should be oriented toward performance. Knowledge or recall should be no more than required when the soldier actually performs the task.

In this respect the written component of the SQT differs in approach from those traditional MOS tests that relied on reading and recalling information that often was not essential to task performance. In fact, the best way for a soldier to "study" for the SQT would be to practice performing the task correctly - not reading about it.

2. TASKS. Tasks for which written scorable units are to be developed will fall into three categories:

- a. Written performance.
- b. Performance-based.
- c. Mixture of written performance and performance-based.

3. WRITTEN PERFORMANCE. A written performance test measures the examinee's ability to perform an actual task at his test station (desk). He performs whatever steps are necessary to arrive at the task product, then chooses his answer from real world alternatives presented in the test booklet.

For example, a written performance SU might involve computing the charge for a 4.2 mortar round. Given certain information and job relevant "tools," the soldier actually would have to perform the task to select the correct answer.

4. PERFORMANCE-BASED. A performance-based test measures the examinee's ability to answer questions about how he would perform a task. For example, an examinee might be asked to identify before operations maintenance checks from a larger list of maintenance checks.

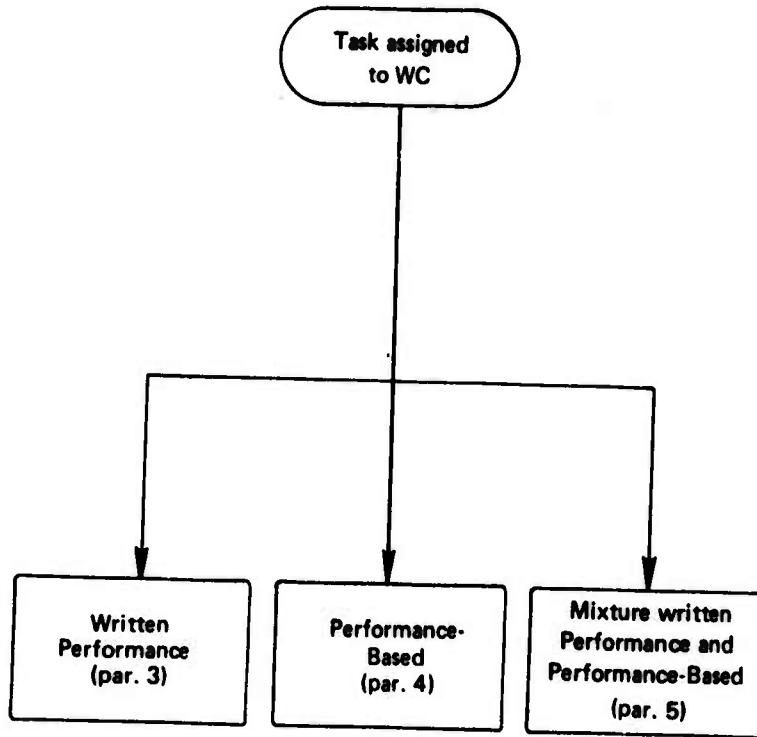


Figure 7-1. Develop Written Scorable Unit Overview

5. MIXTURE OF WRITTEN PERFORMANCE AND PERFORMANCE-BASED TASKS. This category simply combines the two approaches. Many tasks cannot be tested fully using a written performance approach.

For example, given the task "Navigate Cross Country Using a Map," an examinee could be given a mapsheet, protractor, and a situation, and could actually determine coordinates, measure distance and convert grid/magnetic azimuths.

The examinee could then be asked questions about measuring pace and using the compass. We still have not tested his ability to move cross-country, but we have tested his ability to perform several of the steps, and tested his knowledge on how remaining steps should be accomplished.

We can then draw the conclusion that, barring some physical impairment, the individual should be able to "Navigate Cross-Country Using A Map."

6. To determine the type of written SU you will develop, refer to the flow diagram which provides an overview of the written SU procedure. By taking each task assigned to the WC and following the procedure outlined earlier, you will be able to assign each task to one of the three categories.

#### Section I. DETERMINE WRITTEN PERFORMANCE OR PERFORMANCE-BASED

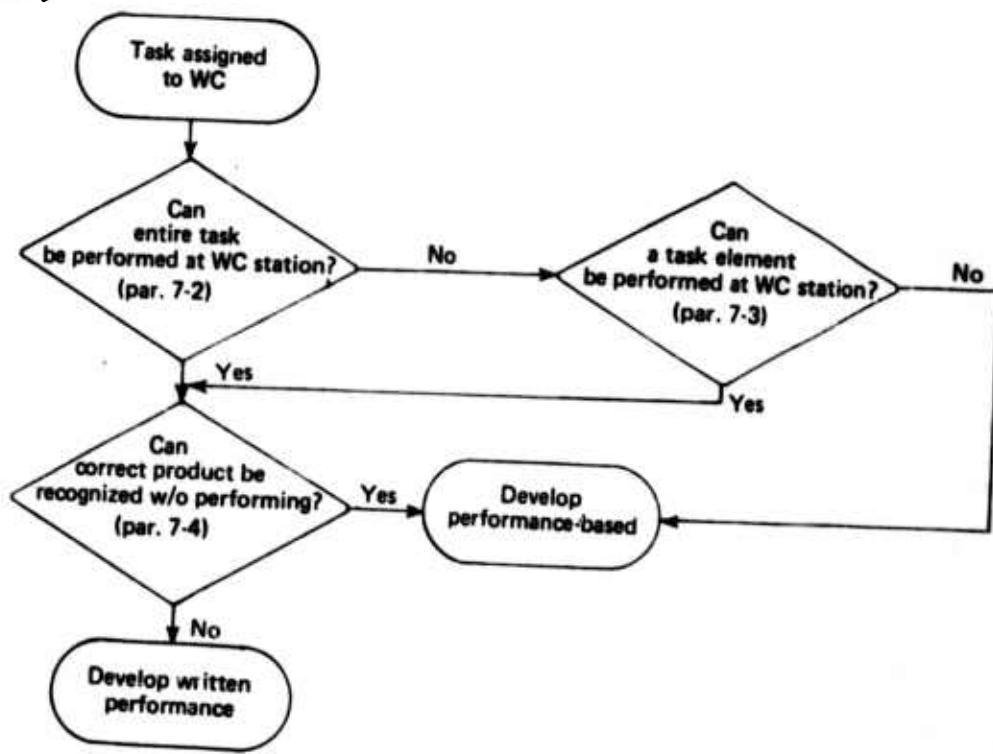
##### 7-1. INTRODUCTION.

a. Generally, if the product of a task can be measured in the written component, you will end up with a written performance SU.

Similarly, if the product of an entire task cannot be tested in the written component, but the product of one or more of the task elements can, at least part of the SU will be written performance.

Where none of the task elements can be product-tested, however, you will end up developing a pure performance-based SU. Depending on the type of task you are working with, many written scorable units will be a mixture of written performance and performance-based items.

The considerations diagrammed in Figure 7-2 should guide your decision as to whether a whole task or task elements can be tested via a written performance item, or whether one or more of the task elements must be tested via performance-based written items.



**Figure 7-2. Determine Written Performance or Performance-Based**

b. A good rule of thumb to use in distinguishing between written performance and performance-based written items is to ask yourself; "Does the item call for the soldier to perform the task or task element as he would perform it on the job, or does the item call for him to answer questions about how he would perform it?"

c. The following explanation should help you answer the questions in the flow chart.

7-2. CAN THE ENTIRE TASK BE PERFORMED AT THE SOLDIER'S WC STATION? In essence, this is asking if the job task can be enacted on the written test, given the limitation of work space (individual work space for the WC is assumed to be equivalent to the area of a standard field table).

Where equipment, tools, space, terrain and so forth are clearly not manageable at the written test station, consider simulation. If, through simulation, you still cannot enact the task within the confines of the written test station, reject the possibility of testing task performance. But it still may be possible to test performance of one or more of the separate task elements.

7-3. CAN A TASK ELEMENT BE PERFORMED AT WC STATION? This question pertaining to an individual task element should be answered in the same way as described above for the entire task. If the answer is "no", prepare a performance-based item for the element. If the answer is "yes", ask yourself the next question.

7-4. CAN CORRECT TASK PRODUCT BE RECOGNIZED BY EXAMINEE WITHOUT ACTUALLY PERFORMING TASK OR TASK ELEMENT?

a. If the correct product of the task or task element can be recognized by the soldier without actually performing the task or element, there is no benefit to having him perform.

Remember, a requirement of the written component is that the soldier must mark his answers on a mark-sense answer sheet. Because of this constraint, the recorded test response made by the soldier will be limited necessarily to a recognition response.

That is, in all cases he will come up with an answer, compare his answer with the listed alternatives, and select the alternative that corresponds most closely to his answer. So, if without actually performing, he can generate a correct answer for comparison with the list of alternatives, it is not a performance test.

Disassembling or assembling a piece of equipment such as an M16 rifle is an example. Having the soldier disassemble the weapon at the written test station accomplishes nothing, since he is able to recognize pictures or verbal descriptions of the product (the weapon in some number of parts) without doing it.

On the other hand, even a skilled soldier could not recognize the correct distance in kilometers along a route on a map without first measuring it.

Therefore, if task product can be recognized without its being produced through performance, you should go to a performance-based testing mode. If the task product cannot be recognized without some performance you have a written performance test.

b. Some examples of task or element products that can be tested in a written performance mode are:

- Calculating the distance along a route on a map.
- Determining range to target using the mil formula.
- Adjusting the M16 sight on the basis of a three-shot group (number of clicks elevation).
- Determining a soldier's eligibility for overseas levy (eligible versus ineligible).
- Determining the minimum readiness status of a piece of equipment (Red, Amber, Green).
- Diagnosing the fault in a piece of equipment (part A versus part B, versus ...).

c. If a task or task element can be performed at a WC station and the correct task product cannot be recognized unless the task is performed, develop a written performance test.

7-5. CONTINUUM. You may find yourself unable to distinguish whether a test item is a performance or is performance-based. Do not be overly concerned with this. While examples can be given of true performance and true performance-based questions, not all test items will fall so neatly into the two categories. Think of the distinction as being along a continuum.

Performance	<u>????????????????</u>	Performance-Based
-------------	-------------------------	-------------------

At either end the distinction is easy. But somewhere in the middle, individual test items become "fuzzy," being neither all one or the other. It is possible to have a good written item that is neither truly performance or performance-based. But to start out with a clear understanding of the differences between the two types of items is still important.

## Section II. DEVELOP WRITTEN PERFORMANCE ITEM

7-6. DEVELOP TEST SITUATION IN WHICH EXAMINEE PREPARES PRODUCT. Whether testing an element or an entire task, begin development by describing the job situation. This statement should provide all information necessary in setting the stage for questions you are going to ask. While necessary information pertaining to task conditions and the job setting must be included, avoid information that is unnecessary to task performance since it consumes valuable testing time.

|See the "General Situation" in the land navigation example  
|(page 7-26, Figure 7-5) which is also an example of an SU  
|that has a mixture of performance and performance-based items.  
|Also included at the end of this chapter is an example of an  
|SU consisting solely of written performance items (Example 2  
|of Figure 7-5).

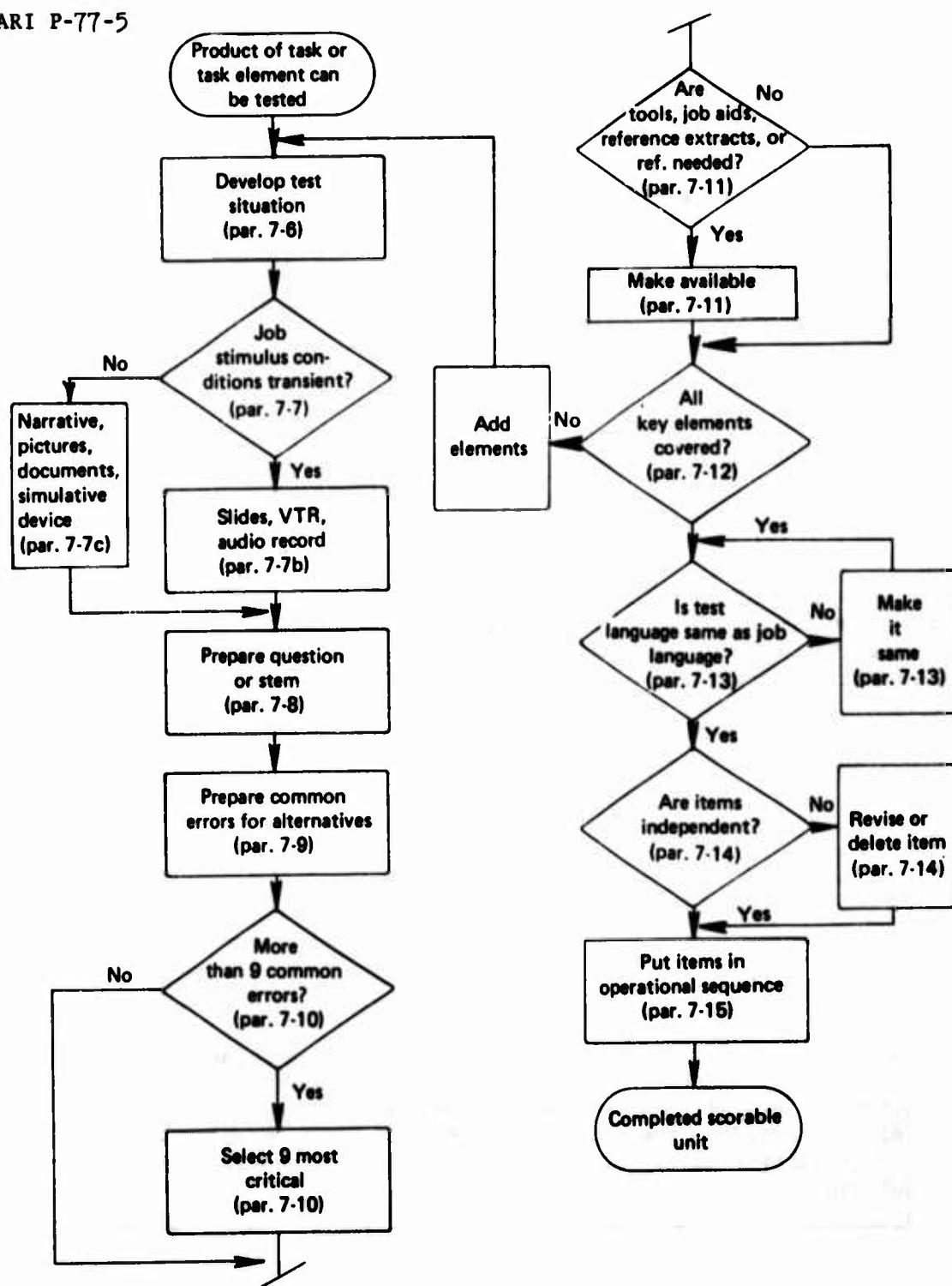


Figure 7-3. Develop Written Performance SU

## 7-7. ARE JOB STIMULUS CONDITIONS TRANSIENT?

a. Some things the soldiers must observe, listen to, or otherwise sense in a job setting may be present for a brief time only in that setting. Examples:

- A passing aircraft that must be identified.
- A briefly exposed tank that must be judged friendly or enemy.
- An incoming radio message to be recorded.

You must recognize when such transient stimuli are part of the task conditions since it will make a difference in how you present the stimulus in the test.

b. Use slides, videotape recordings, or audio recordings. If the stimulus condition is transient, you must attempt to control its presentation to the examinee by use of slides, videotape, or audio recordings.

Although the form of the stimulus will be a simulation (a picture or recorded sound), you can limit its exposure time to that which would occur in the job situation.

c. Use narrative, pictures, documents or simulative device. If a stimulus condition is not transient but relatively constant, for the soldier to react to or operate on, then it should be presented on the test or as a supplement to it. The stimulus can be presented as a narrative description, a picture or diagram, an actual document, or a simulative device. Examples:

- Distance on a map - 25 kilometers.
- Range to a target - 1500 meters.
- Eligibility for overseas levy - yes.
- Readiness status of a piece of equipment - RED.

7-8. PREPARE QUESTIONS OR STEMS. Once you have specified the product of the task, you can prepare a clear question to ask or statement to complete (stem). In testing a task, the question may be preceded by a further narrowing of the general test situation already prepared; that is, a specific situation. The specific situation and question should make it clear to the soldier just what he is to produce.

Examples:

- "What is the ground distance from point A to grid coordinate EG 158858?"
- "After you return from ..., your squad leader gives you a mission to reconnoiter the bridge site at... . Using the grid-magnetic angle shown, what is the magnetic azimuth you will be using?"

7- 9. PREPARE COMMON ERRORS AS ALTERNATIVES. Identify incorrect products that often result in real-world performance of the task. These will be alternatives to the correct answer in the multiple-choice question. No certain number of alternatives should be used in an item.

The number of alternatives is determined by the number of different errors that commonly occur in the real world. Do not simply generate alternatives mechanically, but check to make sure they are reasonable options.

For example, in measuring the distance along a route on a map, the straight line distance (ignoring curves in the route) would be one common error leading to an incorrect product.

Using the scale of miles rather than meters in converting the measurement to meters might be another common error to use in identifying an incorrect answer alternative.

7-10. ARE THERE MORE THAN NINE COMMON ERRORS? Select the nine most critical errors. The SQT answer sheet provides for no more than 10 alternatives. In some cases, there will be no more than 10 real world options (including the correct alternative) that are even possible, much less plausible. These options might be equipment readiness categories: RED, AMBER, GREEN; target classification: friendly vs enemy.

When more than nine common errors occur on the job, reduce them on the basis either or how frequently each occurs in the real world, or how critical each error is. For example, in identifying types of aircraft, there may be more than 10 possible types. Since for a particular question such as "What is the name of this aircraft?" you can list no more than nine alternatives, you might select the nine aircraft most often confused with the one being presented. Or, you could select the nine most critical, in the sense of critical consequences resulting from a particular misidentification.

7-11. ARE TOOLS, AIDS, OR REFERENCES NEEDED? MAKE TOOLS AVAILABLE.

a. If tools, job aids, or references are necessary or considered acceptable in performing the task, they should be made available on the test. You should consider the possibility of using a simulation of the tool or aid if it is not feasible to provide the actual one.

If being able to locate information in a set of references was previously identified as an essential element, manuals or references should be provided.

If it is not feasible to provide the actual references, extracts of the material should be provided to permit an evaluation of that portion of the essential element dealing with the interpretation and use of the information. Example:

• Protractor, compass for map-using.

• TM or extracts from TM.

b. Some cautions are necessary when considering references. It is seldom possible to provide each examinee with a copy of a particular TM or other reference. Written components must be self-contained, i.e., the TCO cannot be tasked to provide references from his resources for the WC.

The most feasible method of providing references is through use of extracts. Up to 30 pages of extracts per test can be printed by ITED. If more than 30 pages are needed, an exception to policy must be requested from ITED along with justification for the exception.

c. For this reason it is probably not feasible to test skills having to do with the ability to locate information in a reference in the WC.

If necessary to test this ability, it should be done in the HOC by developing a special SU for the quasi-task of locating information in a TM or other reference. Limit reference extracts in the WC to situations where the ability to use the information is the critical element.

d. You must rely on your judgment and that of subject matter experts in deciding whether an extract is needed. References cannot be provided for every SU, nor is a reference needed just because it exists. You must consider the job requirements.

For example, the procedure for clearing a misfired round is contained in the TM. Yet the soldier does not refer to the TM when a misfire occurs. You must decide what material the soldier in the MOS and skill level is expected to internalize.

Conversely, do not discard the idea of a reference just because some soldiers in the MOS (usually "experts") can perform the task without use of a reference. This could lead to claims of "foul" in the part of less-skilled examinees who felt discriminated against by lack of references which are authorized and routinely used (by them, anyway) on the job.

Some references, published by commercial industry, developers, or even individual authors, are copyrighted. AR 310-1 specifies the procedure for dealing with copyrighted material and how to obtain clearance for use. If copyrighted material is used, the TDA is responsible for obtaining consent in writing. A copy of the consent should be included with the material when it is submitted to ITED.

7-12. ARE ALL KEY ELEMENTS COVERED? In Chapter 3 you identified key elements (those steps the soldier is most likely to fail or steps that have serious consequences of failure). Now, be sure that all identified key elements are covered in the written scorable unit.

Other steps or elements may be covered also, but there must be at least one test item for each key element. If any are not covered, go through this development procedure for each remaining element. It may be necessary to use the performance-based model in Section III to insure all key elements have been covered.

7-13. IS TEST LANGUAGE SAME AS JOB LANGUAGE? Be certain that the test items and test situation communicate to the examinee. In more technical areas, a tendency is to try to simplify technical items during development of items. When done, however, the item may lose its intended meaning.

Technical terms used on the job by the examinee have very precise meanings. If used routinely on the job, the terms probably should not be changed or "simplified" for the test. All test materials should be written in the normal job language and reviewed and revised until they pass this checkpoint.

7-14. ARE ITEMS INDEPENDENT? This decision point is included to insure that the stem for one item does not give away the answer to another item. An examinee who reads the stem to an item should not be able to work backwards and correct a previous answer.

Check to make sure you have not answered any items in the test situation or the item stems. If you cannot reword or restructure a cueing item, delete the least important item that provides cueing.

7-15. PUT ITEMS IN OPERATIONAL SEQUENCE. The final action of this process is to place items in a scorable unit in operational sequence. In developing this sequence, list items in the order that they normally would occur on the job.

### Section III. DEVELOP PERFORMANCE-BASED WRITTEN SCORABLE UNIT

#### 7-16. INTRODUCTION.

a. The performance-based written SU likely may be the most difficult to develop. Yet, because of the limitations on HOC and the relatively limited number of items that can be developed in the written performance mode, it is the type of SU that will be used most frequently for SQT.

Development is difficult because we are still trying to test performance, not abstract knowledge or recall. Since the test is performance-based, direct your attention to what happens when the soldier performs the task. Specifically, focus on the reasons that soldiers frequently cannot perform the task.

b. As illustrated in the flow diagram, five main categories of reasons show why a soldier may fail to perform a task. Some tasks will fit more than one of the categories.

When going through the flow diagram, go through each of the five top decision diamonds before going through the method below each. This is to insure that all categories have been considered in selecting the most appropriate causes of failure.

#### 7-17. WHY IS PROFICIENCY LOW? Each element or step in the task must be considered.

The developer should utilize his own knowledge of soldiers in the skill level, available subject matter experts, plus any information he has gathered during the task selection process to determine most likely reasons a soldier cannot perform the step.

Although no formal evaluation of the step is required, use the best resources you have available in making this decision. Take each element through the following five decision points. However, not all elements need be included in the test. Some elements simply pose no performance problems for soldiers and these should not be included in testing.

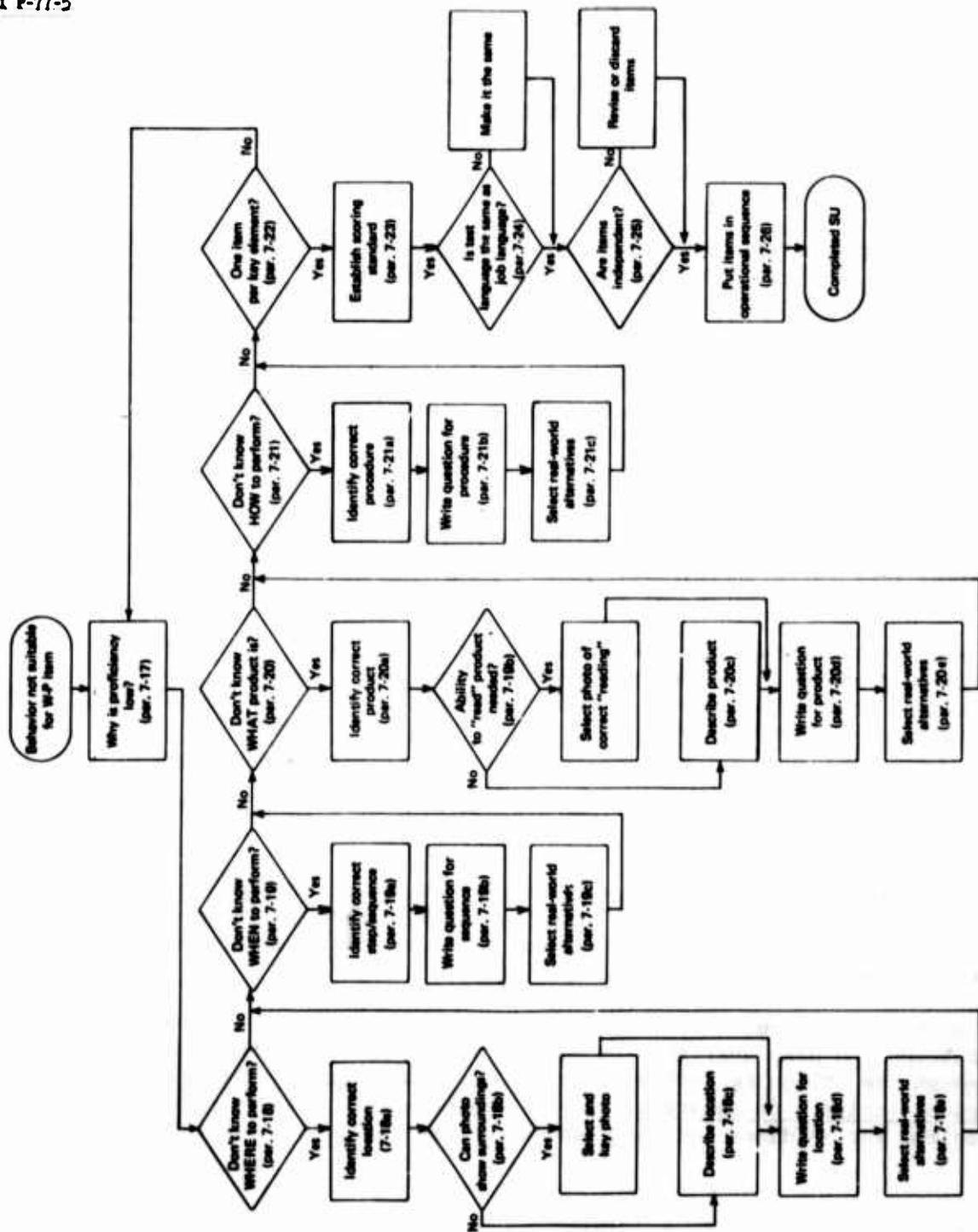


Figure 7-4. Develop Performance-Based Items

7-18. DO NOT KNOW WHERE TO PERFORM.

The failure to perform some tasks may be due to the inability to locate certain objects. For example, a vehicle driver may fail to perform checks on preventive maintenance on the component of a vehicle because he does not know where that component is located.

A mechanic may fail to zero a multimeter because he does not know the location of the correct adjustment screw. If location is the problem, follow the procedure below. (If location is not the problem, go to Paragraph 7-19.)

a. Identify the Correct Location. This is simply deciding what the right answer is. It may be necessary actually to view the equipment if the element is equipment-oriented. Technical manuals may not have a clear picture or a clear description. Be alert for TMs which may be in error.

b. Can Photographs Show Item and Surroundings? It may be possible to show locations by the use of photographs. In fact, clear photos approximate the job presentation closer than any other two-dimensional means. If photos are not practical, consider other visual means such as drawings or schematics. Insure that the photo or drawing is properly keyed.

c. Produce Description of Location. If surroundings of the component cannot be illustrated clearly and accurately in a picture or visual representation, then location of the component should be described as clearly as possible.

d. Write Question(s) for Location. Develop one or more questions which will examine thoroughly reasons for failure to locate the component. For example, with a keyed picture of a multimeter, the question might be:

"Which letter in Figure \_\_\_ represents the location of the zero adjusting screw?" Or if no picture, the question might be: "Which of the following best describes the location of the zero adjusting screw on the multimeter?"

e. Select Real World Alternatives. Consider the likely options which you would find on the job. Select from or use these options (no more than nine since there are only ten responses per question) as alternatives to the correct answer.

The types of alternatives selected will vary depending on whether the item uses visual or verbal information. For a given item, the number of alternatives selected is not important. It is important, however, that the alternatives be reasonable and realistic in terms of errors that commonly occur in the real world.

7-19. DON'T KNOW WHEN TO PERFORM STEP. The next major reason for failure to perform a task is that the soldier may not know when in a sequence of activities the step should be performed. For example, it is physically possible to remove the hand guards on the M16 rifle at almost any time during assembly.

However, there is a point in the sequence when this step should be performed if the soldier is to avoid damaging the weapon. If the problem is determined to be one of when or sequence, use the following procedure. (If not a problem of when, go to paragraph 7-20.)

a. Identify Correct Step/Sequence. This means identifying the correct point in the sequence for performing the step.

b. Write Question for Sequence. Develop one or more questions that will thoroughly examine the reasons for failing to perform this step in sequence. The question might be phrased this way:

"Which of the following steps must be performed before ...?" or, "Which of the following steps must not be performed until ...?" If the problem is one of soldiers performing the step, but at the wrong point in the task, the step may be given either in the question or as one of the alternatives.

But if the problem is one of soldiers leaving the step out, it would be better to offer it as an alternative while referring to a preceding (or succeeding) step to key the sequence in the question.

c. Select Real World Alternatives. Identify reasonable alternatives that the soldier would find on the job. Select alternatives with emphasis on those variations which include adjacent steps. Remember that nine alternatives (plus the correct answer) are the limit. Only include the likely variations if there are more than nine.

#### 7-20. DON'T KNOW WHAT THE PRODUCT IS.

Another reason for failure to perform a task correctly is that the individual may not know what the end result of the element or step should look like.

For example, a clerk cannot type a correct address label unless he knows how a proper address is formatted. A gunner cannot obtain a proper sight picture unless he recognizes the sight alignment when he sees it. If this is the cause of soldiers' failure, use the following procedures. (If not the cause, go to paragraph 7-21.)

a. Identify Correct Product. The first step is to identify correctly what is the correct outcome, result, or product of the step.

b. Is Ability to "Read" Product Needed? In determining whether the correct product has been obtained, an individual may be required to judge or interpret diagrams, meter readings, aerial photographs, charts, position settings, or other similar visual representations.

If this kind of thing is used in determining whether or not the step has been completed, then pictures or visual representations should be used in the test items.

c. Prepare Description of Product. If the product cannot be represented by some visual means, it is necessary to write out a description of the product. Insure that the description contains the discriminating cues that the soldier mentally uses on the job.

For example, "needle reading is in the green area" or "the dial is rotated until the detent is engaged."

d. Write Question(s) for Product. Develop one or more questions which will examine thoroughly the reasons for failing to produce (identify) the correct product. These questions should center around the scope and basic content of what the correct product is.

For example, "Which of the following pictures shows the correct sight alignment for the M16?" or, "When conducting an operational check of the field telephone, the LOUD switch should be in which of the following positions?"

e. Select Real World Alternatives. Identify all of the reasonable variations and errors expected to occur on the job. Select alternatives from these common errors which would lead to failure to produce/identify the correct product.

Alternatives selected should reflect the more critical errors if more than nine are identified.

#### 7-21. DON'T KNOW HOW TO PERFORM PROCEDURE.

The final consideration for why soldiers cannot perform has to do with procedure. Soldiers may know where and when to perform a step, and what the correct result is, but they may not know how to produce that result - that is, how to execute the step.

Procedural failures may result because the soldier simply does not know what the procedure is or because he forgets it.

For example, in preparing the light antitank weapon (LAW) for firing, one step is to extend the launcher tube. The soldier may know that the tube must be extended prior to firing and where the tube is located and what the LAW would look like with the tube extended. But unless he knows how to extend the tube, he cannot perform the task. If your element falls within procedural reasons for failure, follow the steps below.

a. Identify Correct Procedure. As with all reasons for failure the first step is to identify the correct way to perform the element.

b. Write Questions for Procedure. Develop one or more questions which will examine the reasons for failing to perform the procedure correctly. Several approaches are possible. The soldier may be asked to select correct steps from several presented, to place steps in sequence or to judge a series of photographs as to their correctness.

c. Select Real World Alternatives. Select the nine (or less) responses as alternatives that reflect the most likely incorrect actions that the soldier would take on the job.

7-22. IS THERE ONE ITEM FOR EACH KEY ELEMENT? Recall the earlier work (Chapter 3) in selecting key elements. Now insure these key elements are covered by test items. You may decide to test elements not designated as key elements as well, but the key elements must be covered.

Each key element should be evaluated by at least one item to provide a comprehensive examination. In some cases it may not be realistic to develop more than one item for each key element.

Examine test items for each key element and determine if critical aspects of the element have been omitted in terms of location, sequence, product, or procedure. Develop additional test items as required to provide a comprehensive evaluation of each key element.

7-23. ESTABLISH A SCORING STANDARD.

a. Once you have developed a Written SU, you must set a passing score. This involves deciding the number of items the examinee must answer correctly in order to pass the SU. In most cases this is all you will have to do.

Sometimes, however, your work will be complicated by an item in the SU having more than one correct answer. Then you will have to decide on a passing score for the item also. The following offers guidance on determining a passing score for an SU and for an item with multiple correct answers.

**b. Passing Score for Scorable Unit.**

(1) Establishing a passing score for an SU is not as straightforward as we might think. One can argue convincingly that, since we are dealing with performance-oriented criterion-referenced testing, the soldier must pass every item in order to pass the SU. And in theory that is true.

There are however extenuating circumstances - particularly when written tests of task performance are used. Words sometimes get in the way of a soldier trying to communicate his or her knowledge of task performance.

Differences in ability to read and understand the printed word or to focus mental concentration on words for a prolonged time may cause examinees to misread or misinterpret occasionally some parts of the written test, no matter how carefully the developer worded the questions.

Since these lapses should not necessarily be interpreted as a lack of knowledge about how to perform the task, we should allow the examinee some small margin of error in the Written Component.

Research has shown, in fact, that written test scores are truer (reflect ability actually to perform) when some margin of error is allowed.

(2) Recommended passing scores for SU's of various sizes are given below. Note that the examinee is not allowed an incorrect answer for SU's with 4 or fewer items; but is allowed one miss for an SU with 5 to 7 items, and two misses for one with 8 to 10 items. Using these standards for scoring SU's should increase the validity and perceived fairness of your Written Component.

Number of Items Examinee  
Must Answer Correct to Pass an SU

c. Passing Score for Item.

(1) For reasons similar to those just given, we should consider also allowing the examinee some margin of error on items that have more than one correct answer. Such items tap several knowledge elements at once, elements that would be covered separately in a more conventional multiple-choice test.

But since with SQT questions, we are striving to maintain task integrity, you may develop items that call for more than one correct answer. Those items tend to be more difficult for examinees, especially when they are not told how many alternatives to mark.

(2) The conventional four-alternative multiple-choice item has one correct answer. The examinee, even when in doubt, marks but one alternative, knowing that he will surely fail the item if he marks more than one or none at all.

In contrast, an examinee responding to an item with multiple correct answers, if at all uncertain about one of the alternatives, must decide how many to mark while deciding which ones are correct; that is, not being told the exact number to mark makes it difficult to know whether some questionable alternative should be included.

For this reason, items with multiple correct responses require more precise knowledge of the subject matter, and the examinee should be allowed some small margin of error.

(3) Below is a table of recommended passing scores for items with multiple correct answers. Notice that to make the table complete, we have included the case of one correct response. More importantly, notice that the passing scores shown in the body of the table are based on all alternatives being scored.

The maximum error allowed tends to be where about half of the alternatives are keyed as correct. In no case, however, is more than one error each of omission and commission allowed.

**Recommended Passing Scores\* For  
Items With Multiple Correct Answers**

	Number of Item Alternatives							
	4	5	6	7	8	9	10	
1	4	5	6	7	8	9	10	
2	3	4	5	6	7	8	9	
3	4	4	5	5	6	7	8	
4		5	5	5	6	7	8	
5			6	6	6	7	8	
6				7	7	7	8	
7					8	8	8	
8						9	9	
9							10	

\* The passing scores shown in the table are based on all alternatives, i.e., both marked and unmarked alternatives scored.

(4) To use the table, find the column representing the number of alternatives in your test item, then read down that column to the row representing the number of correct alternatives.

The number at the intersection of that row and column is the minimum number of item alternatives that must be correctly marked (and left unmarked) for the examinee to pass the item.

For example, suppose you have an item with eight alternatives, four of which are correct. The recommended passing score is six. A soldier who marked two or three of the four correctly and did not mark any of the incorrect alternatives, would pass.

If he marked all four correct alternatives, plus no more than two of the incorrect ones, he would also pass. Or if he failed to mark one of the correct ones but marked no more than one of the incorrect, again he would pass.

(5) If these passing standards seem too lenient, remember that in no instance do they allow more than one combined error of omission and commission.

7-24. IS THE TEST LANGUAGE THE SAME AS THE JOB LANGUAGE? You must insure that the test items and test situation communicate to the examinee. In some of the more technical areas, there is a tendency to try to simplify technical terms during the development of items ostensibly in the interest of communication.

When this is done, however, the item may lose its intended meaning. Technical terms are used on the job because they have very precise meanings. All test materials should be written in the normal job language and reviewed and revised until they conform completely with the current job language.

7-25. ARE ITEMS INDEPENDENT? This decision point is included to insure that the answer to one question does not influence or is not used in the selection of the answer to another question. This is not to say that the items cannot be related.

The test situation for an SU normally will provide a relationship between all the items in that unit, but the answer to one question should be completely independent of the selection of the answer to another question.

Items should be independent from the standpoint of cueing. The stem of an item should not cue the answer to a previous question. If test items are found that are not independent, the items should be revised.

7-26. PUT ITEMS IN OPERATIONAL SEQUENCE. The final activity of the process is to place the items in an SU in operational sequence. In developing this sequence, list the items in the order that they would occur on the job.

**Example 1 - Determine Parachute Release-Weight Setting for Equipment Drop**

You have just completed rigging at 19,500 lb. load for drop by six G11-A parachutes. The parachutes are rigged with 5,000 lb. release assemblies. What is the release weight?

- |              |             |
|--------------|-------------|
| 1. 18,000 lb | 5. 3,000 lb |
| 2. 14,500 lb | 6. 2,500 lb |
| 3. 3,900 lb  | 7. 2,000 lb |
| 4. 3,500 lb  | 8. 1,500 lb |

**EXAMPLE: Performance-Based Scorable Unit****TASK: Set Up Field Telephone and Contact Another Station****ESSENTIAL BEHAVIORS:**

Strip 1" insulation from wire

Connect field wire to binding post

Install batteries (one + up, other + down)

\*Turn circuit selector switch to LB

Turn buzzer volume control to LOUD

Turn EXT-INT switch to INT

\*Signal station by turning crank with handset in retaining cradle

\*Key Behaviors

**Figure 7-5. Examples of Written-Performance Items (page 1 of 6).**

Your weapon system is in a fixed defense of an airfield. Communication is by field telephone. Field wire has been laid from Switchboard SB-993/GT to your position. The Telephone Set, TA-312/PT has just been issued from supply. The power source will be two BA-30 batteries.

1. How would you prepare the field wire?
  - a. Strip last 3 inches
  - \*b. Strip last inch.
  - c. Leave insulation intact.
2. How would you install the two batteries in the TA-312/PT?
  - \*a. One with positive (electrode) end up, the other with negative (smooth) end up.
  - b. Both with positive (electrode) end up.
  - c. Both with negative (smooth) end up.
3. Where on the TA-312PT would you secure field wire?
  - \*a. Binding posts.
  - b. External battery terminals.
4. Which is the correct position for the circuit selector switch?
  - a. CBS
  - b. CB
  - \*c. LB
5. Which is the correct position for buzzer volume control knob?
  - a. LOW
  - b. Halfway between LOW and LOUD
  - \*c. LOUD
6. What is the correct position for the EXT-INT switch?
  - a. EXT
  - \*b. INT
7. How would you signal another station?
  - a. Press push-to-talk switch.
  - \*b. Turn generator crank with handset securely in retaining cradle.
  - c. Turn generator crank with handset removed from retaining cradle.

\* Key behaviors

Figure 7-5. (page 2 of 6)

## Example 2 - Determine Distance Along a Route on a Map

You are at point A (See map below, Figure 7-5) and have been ordered to move your squad to a position at B. Since the road is secure, you plan to move along that route. What is the distance in meters along the road from A to B.

- |          |          |           |
|----------|----------|-----------|
| 1. 1,900 | 5. 3,000 | 9. 3,400  |
| 2. 2,300 | 6. 3,100 | 10. 3,500 |
| 3. 2,700 | 7. 3,200 |           |
| 4. 2,900 | 8. 3,300 |           |

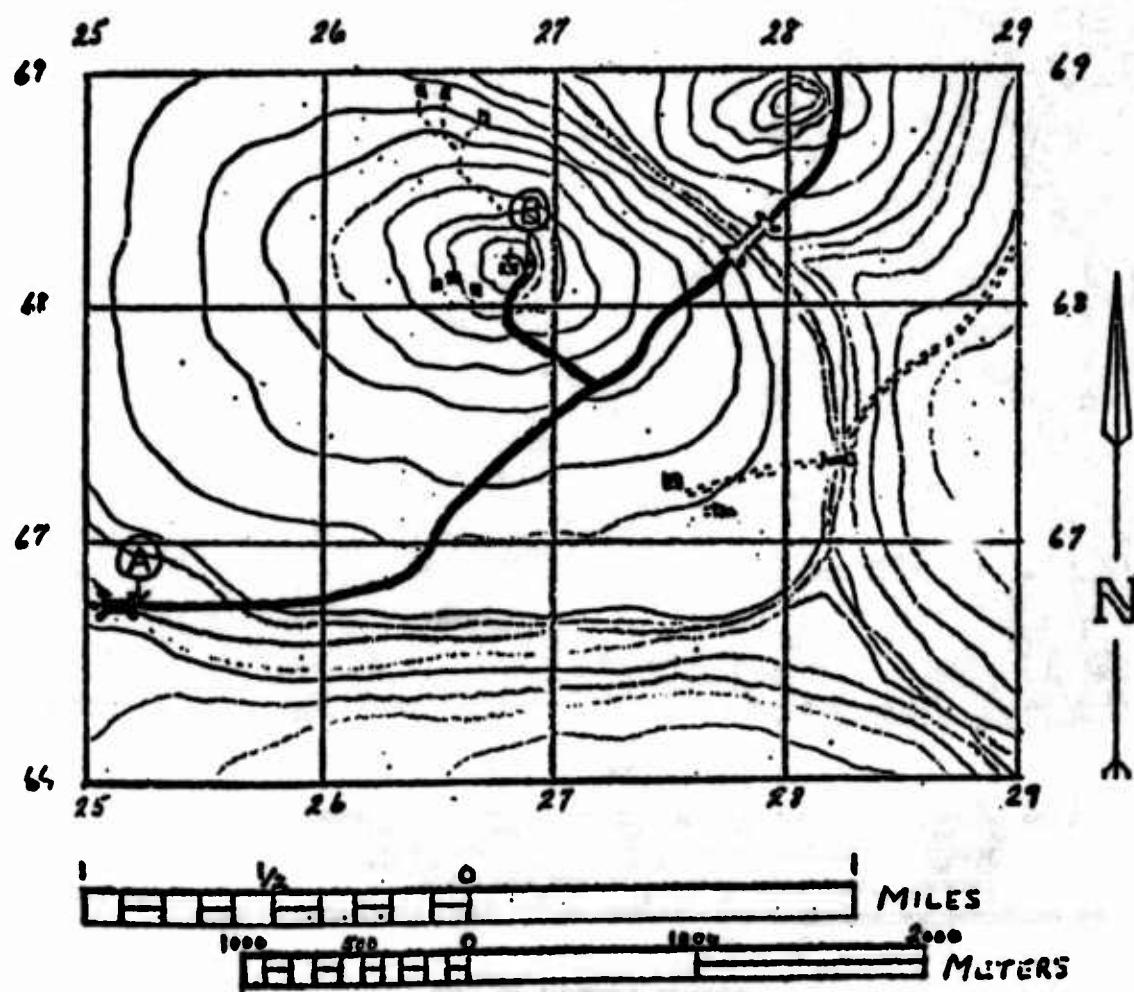


Figure 7-5. (page 3 of 6)

ARI P-77-5

EXAMPLE: SU with a Mix of Written Performance and Performance-Based Items

Title: Navigate from one position on the ground to another point.

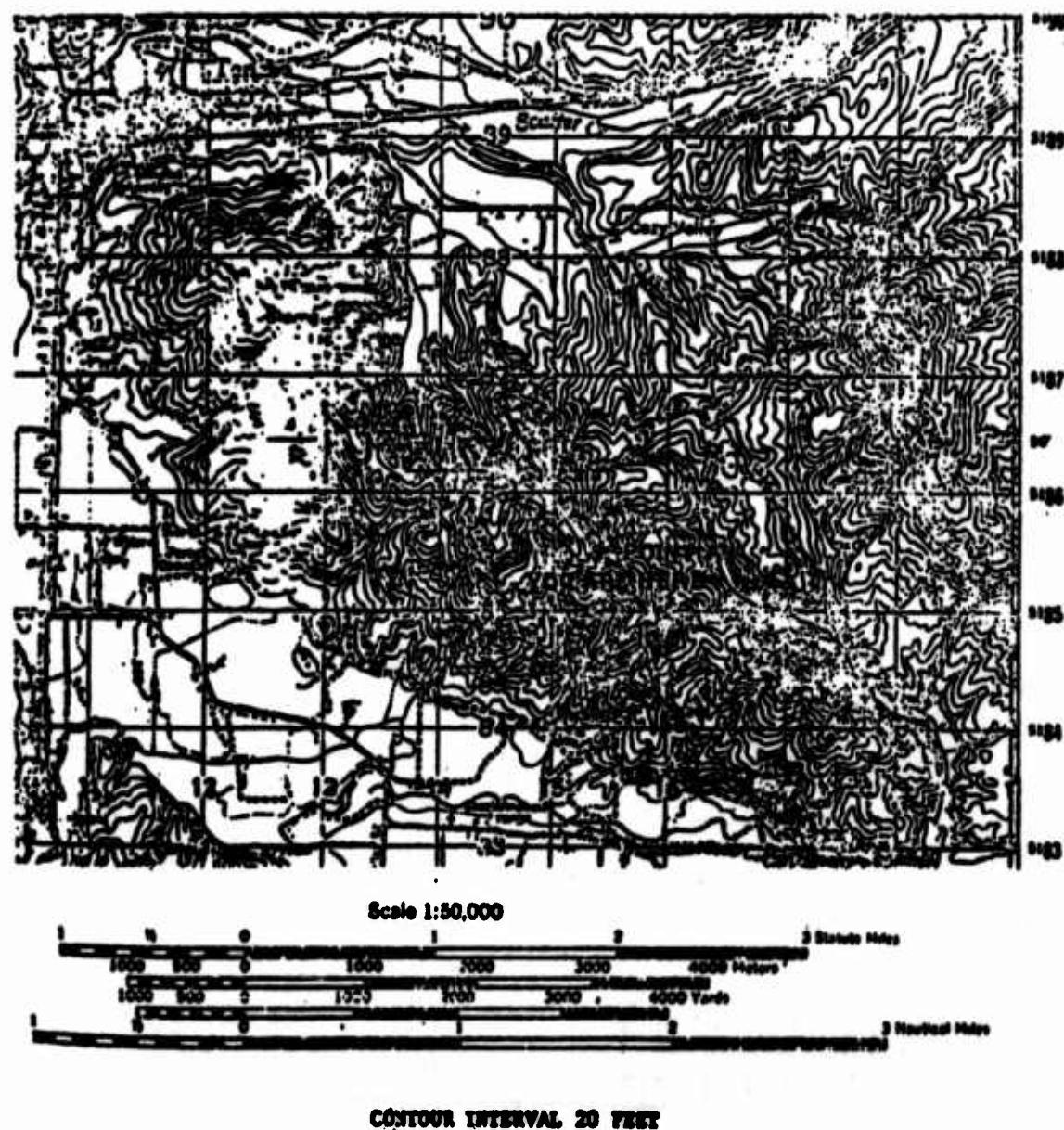


Figure 7-5. (page 4 of 6)

General Situation: You are a fire team leader. Your squad leader has just told you to take your fire team to the company trains (supply point) to pick up rations and ammunition for the platoon. He hands you a map and tells you that you are at Point A. He wants you to go to grid coordinate EG 158858. (A protractor/grid coordinate scale is provided in each test packet.)

1. Special Situation: Using the map above, what is the six-digit coordinate of Point "A"?

- a. EG 155856
- \*b. EG 155854
- c. EG 154857
- d. EG 153856
- e. EG 153854

2. Situation continued: What terrain feature is located at grid coordinate EG 158858 (where you are going)?

- a. Hilltop
- b. Ridge
- \*c. Valley
- d. Saddle
- e. Depression

3. Situation continued: What is the ground distance from Point "A" to grid coordinate EG 158858?

- a. 350 meters
- b. 450 meters
- \*c. 500 meters
- d. 650 meters

4. Situation continued: In this case the grid-magnetic (GM) angle is so small that the grid azimuth and the magnetic azimuth is the same. What magnetic azimuth should be followed going to grid coordinate EG 158858?

- |                |               |
|----------------|---------------|
| a. $28^\circ$  | d. $43^\circ$ |
| *b. $33^\circ$ | e. $48^\circ$ |
| c. $38^\circ$  | f. $53^\circ$ |

5. Situation continued: You are at the company supply point packing up your platoon's supplies and the first sergeant tells you your platoon has just moved to grid coordinate 171843 (Point B). What is the grid azimuth to your platoon's new location?

- \*a. 125
- b. 140
- c. 305
- d. 320

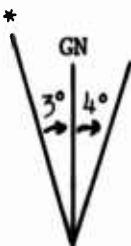
Figure 7-5. (page 5 of 6)

ARI P-77-5

6. Situation continued: What is the elevation of the new platoon position? (The contour interval is 20 feet.)

- a. 720
- b. 880
- c. 900
- d. 920

7. Situation continued: After your return from the company supply point, your squad leader gives you a mission to reconnoiter a bridge site located at map coordinate 195855. The grid (map) azimuth to the bridge from your location is 47. Since you will be using a lensatic compass you must convert the grid (map) azimuth to a magnetic (compass) azimuth. Using the grid-magnetic angle shown below, what is the magnetic azimuth you will be using?



- a. 40°
- \*b. 43°
- c. 51°
- d. 54°

8. If you travel 100 meters in 130 paces, how many paces would you take to travel 800 meters?

- a. 800
- b. 840
- c. 940
- \*d. 1040

9. Where should the black index line on the lensatic compass be set when you navigate cross country?

- a. Directly over the North arrow
- b. Directly over the grid azimuth
- \*c. Directly over the magnetic azimuth
- d. Directly over the grid azimuth divided by 3
- e. Directly over the magnetic azimuth divided by 3

Figure 7-5. (page 6 of 6)

## CHAPTER 8

### VALIDATE WRITTEN COMPONENT

#### Section I. TRY OUT WRITTEN SU

8-1. INTRODUCTION. The tryout is a systematic way to have experts in the MOS review the WC to be sure each SU is consistent with doctrine. You must test experts and collect their opinions of the SU before submitting the WC to ITED.

8-2. ADMINISTER SU TO TEST DEVELOPERS. Before you begin testing experts, two or three people who helped develop an SQT should take each SU. Test developers do not have to be qualified in the MOS covered by the SQT.

8-3. IS THE SU ACCEPTABLE? Ask each test developer to suggest changes in the SU. Test developers who take the test should look for cues within each item or in other items, check readability of items and check clarity of illustrations.

8-4. REVISE THE SU. Weigh suggestions of other test developers. If a suggestion makes sense, change the item or items. Have test developers review the revised SU.

#### 8-5. SELECT AT LEAST THREE EXPERTS.

a. In this case "experts" are people who you are sure can do the task covered by an SU. They do not have to be in the skill level for the SQT. In fact, experts usually will be senior NCO's or instructors in the MOS. One person may be an expert for several tasks covered by an SQT. Test developers should not act as experts for this part of the tryout.

b. The quality of the tryout depends on the level of expertise of the soldiers who take the SU. Do whatever you can to get high-level performers for the tryout. If you have doubt about qualifications of a potential expert, interview his peers and supervisor. Ask the peers and supervisor how well the expected expert can perform the specific task. If they say the expected expert can do the task very well, consider including him in the tryout. If they say the expected expert can do the task only fairly well or worse, do not include him.

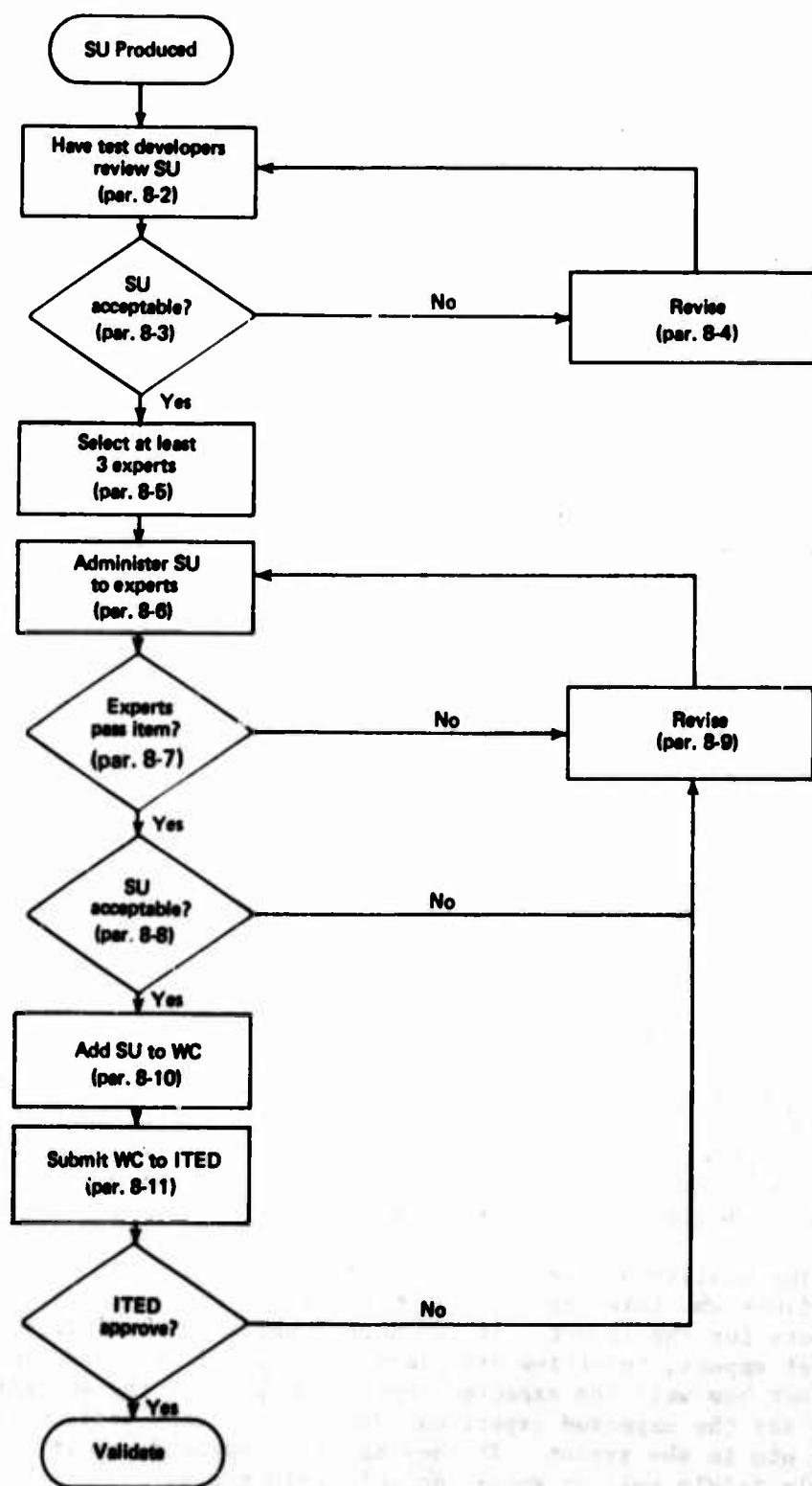


Figure 8-1. Try Out Written Component SU

c. After you locate several expected experts, give them a copy of the task analysis and have them rate their ability to perform the task. When collecting self-ratings, emphasize importance of accuracy. Be sure expected experts know there is no benefit to them or the SQT if they overrate their ability. Use one of these questions for self-ratings. The question to choose depends on whether the task is performed normally under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well.
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

For this tryout, experts are people who choose No. 3, "Very well" or "No guidance."

8-6. ADMINISTER SU TO EXPERTS. Test experts one at a time. Make conditions as similar as possible to field conditions. Use the same instructions and illustrations as when testing for record. Allow time for the expert to answer items and review answers.

8-7. EXPERTS PASS EACH ITEM? One minimum requirement of an SQT-written SU is that experts on a task consistently pass items based on that task. If an expert misses any item, find out why. The expert may have forgotten how to do the task or the task may have changed. Otherwise, revise the item to eliminate the characteristic that led to the incorrect answer.

8-8. IS THE SU ACCEPTABLE TO EXPERTS? Have the experts complete the questionnaire in Figure 8-2 for each SU immediately after they finish the SU. The questions ask experts to evaluate criticality of the task, fairness of the items, and inclusiveness of the SU. If more than one-third of the experts think the task is non-critical, an item is unfair or the SU is incomplete, revise the SU. Questions relevant to this decision are marked with an asterisk (\*).

8-9. REVISE THE SU. Your primary concern during this tryout is that the SU is consistent with doctrine. If experts select an incorrect alternative, check the alternative to be sure it is a real-world alternative and correct. If experts think the task is non-critical, consider dropping it from the component. If experts think an important behavior has been left out of the SU, develop an item for that behavior. Have the experts review the revised SU.

8-10. ADD THE SU TO THE WC. A written SU that experts consistently pass and consider acceptable meets first requirements for inclusion in the SQT. Add it to the component and repeat the procedure for each remaining SU.

8-11. SUBMIT WC TO ITED. Submit the following things to ITED:

- All SU in the WC.
- A summary of the tryout.
- The task analysis data for each task covered.
- Administrative instructions.
- The SQT notice for the WC.
- Choice of two SU to be validated against hands-on (Figure 8-4).

Submit the WC at the same time you submit the HOC.

Once the WC has been approved by ITED the content of each item within each SU is fixed and cannot be changed without approval from ITED. Items found to be unsatisfactory during validation will be revised to improve them, but not to extent of changing content. Unsatisfactory items cannot be deleted without approval.

**INSTRUCTIONS:** Please take attached scorable unit and answer questions to the best of your ability. Now look at this task in the applicable Soldier's Manual and answer the following questions:

**PART I - Task Analysis**

\*1. Is this task a critical task for the MOS and skill level for which it is written? If no, explain.

- a. Yes                  b. No                  c. Not sure

\*2. Have any steps (performance measures) been left out of the Soldier's Manual? If yes, explain.

- a. Yes                  b. No                  c. Not sure

\*3. Have any unnecessary steps (performance measures) been included? If yes, explain.

- a. Yes                  b. No                  c. Not sure

**Figure 8-2. Subject-Matter-Expert Questionnaire (To be used in review of written scorable unit) (page 1 of 5)**

PART II: Written Scorable Unit Evaluation

INSTRUCTIONS: Now, review the scorable unit and answer the following questions:

4. How many questions (performance measures) should a soldier answer correctly in order to be given a GO on the task:

\_\_\_\_\_ of \_\_\_\_\_

5. Do you think you received a GO (pass) or NO-GO (fail) on the unit?

- a. Go                    b. NO-GO                    c. Not sure

\*6. Is this unit realistic? Circle one. If no, state why.

- a. Yes                    b. No                    c. Not sure

\*7. Is the unit doctrinally correct? If no, state why.

- a. Yes                    b. No                    c. Not sure

8. Would an illustration or additional illustrations make the unit (or any specific item) easier to understand? (If yes, what illustration(s) should be added, and where?)

- a. Yes                    b. No                    c. Not sure

Figure 8-2. (page 2 of 5)

\*9. If the unit has an illustration or illustrations, was it (or they) easy to understand? If no, tell why.

- a. Yes                    b. No                    c. Not sure

\*10. Were words or phrases in the unit difficult to understand? If yes, which ones?

- a. Yes                    b. No                    c. Not sure

\*11. Did you understand what you were asked to do in each item of the unit? If no, tell why.

- a. Yes                    b. No                    c. Not sure

\*12. Do you agree with the correct solution(s)? If no, tell why.

- a. Yes                    b. No                    c. Not sure

13. In the following space, tell which errors a soldier most often makes in performing the task on which this scorable unit is based.

\*14. Review the answers to each question. Are all of the answers possible "real-world" alternatives? If no, which answers are unrealistic?

- a. Yes                    b. No                    c. Not sure

Figure 8-2. (page 3 of 5)

**ARI P-77-5**

**INSTRUCTIONS:** These questions apply to your overall evaluation of the scorable unit.

\*15. Should questions (items) be added to the scorable unit? If yes, explain.

a. Yes

b. No

c. Not sure

\*16. Should questions (items) be deleted from the scorable unit? If yes, explain.

a. Yes

b. No

c. Not sure

\*17. As the scorable unit is now written, do you think it is fair measure of a soldier's ability to perform the task on which it is based? If no, explain.

a. Yes

b. No

c. Not sure

18. If your recommendations are included in a revision of the scorable unit, will it then be a fair measure of a soldier's ability to perform the task on which it is based? If "no", explain.

a. Yes

b. No

c. Not sure

**Figure 8-3. (page 4 of 5)**

INSTRUCTIONS: Now read the portion of the Soldier's Manual in which this scorable unit is based and answer the following questions.

19. Can a soldier answer questions in the scorable unit just by studying the applicable Soldier's Manual?

- a. Yes                    b. No                    c. Not sure

20. (If your answer is yes, to question 19, skip this question). If you answered no or not sure, look at the reference materials. Can the soldier answer questions in the scorable unit correctly by studying or using referenced materials (FMs, TMs, AR's etc.)? If no, tell why.

- a. Yes                    b. No                    c. Not sure

21. Are task, conditions, and standards tested in the scorable unit the same as task conditions and standards in the applicable Soldier's Manual? If no, tell why.

- a. Yes                    b. No                    c. Not sure

INSTRUCTIONS: If applicable, review the TEC Lesson(s) which cover the task on which this scorable unit is based and answer the following questions.

22. Are the task conditions and standards in this scorable unit the same as the task, conditions and standards in the applicable TEC lesson? If no, tell why.

- a. Yes                    b. No                    c. Not sure

23. Would the TEC lesson be of help to a soldier preparing for the SQT if this scorable unit is on the SQT? If no, tell why.

- a. Yes                    b. No                    c. Not sure

Figure 8 3. (page 5 of 5)

ARI-P-77-5

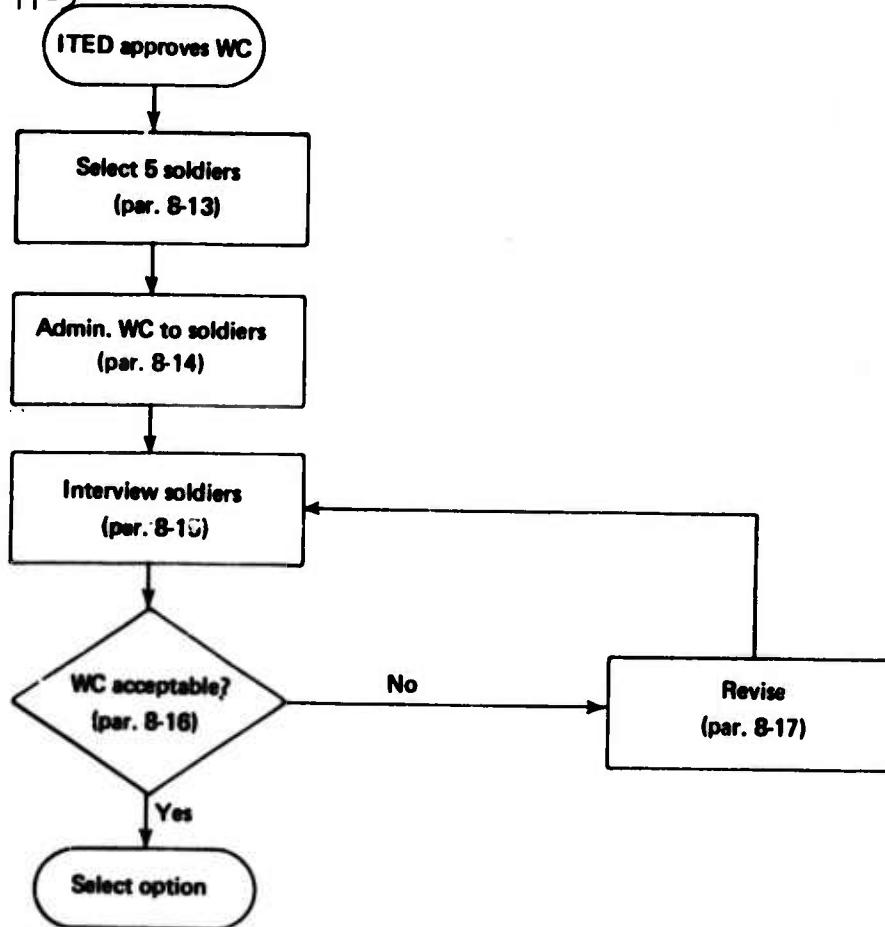


Figure 8-3. Try Out Written Component

Section II. TRY OUT WRITTEN COMPONENT

8-12. INTRODUCTION. This tryout will be first chance to find whether the whole component is appropriate to a sample of representative soldiers in the MOS.

8-13. SELECT FIVE SOLDIERS. After ITED approves the WC, conduct an informal tryout of the component. The first step is to locate five soldiers at the skill level of the MOS covered by the SQT.

8-14. ADMINISTER WC TO SOLDIERS. Test soldiers on the SU that are critical to their skill level. Make conditions as similar as possible to the field conditions. Use the same illustrations and instructions as when testing for record. Allow soldiers time to answer each question and to review their answers.

8-15. INTERVIEW SOLDIERS. After soldiers finish the component, ask each one if instructions or items were hard to understand.

8-16. IS WC ACCEPTABLE? This informal tryout should answer two questions:

- Can soldiers at the skill level understand the instructions, questions and alternatives on the test?
- Can soldiers at the skill level finish the test in the time allowed? Since this tryout does not include items for the next highest skill level, anything over three hours should alert you to a possible problem.

8-17. REVISE. Make only minor revisions based on this tryout. Add illustrations or reword confusing items; do not change content of any SU. Have soldiers in sample check revisions to assure you have eliminated the source of confusion.

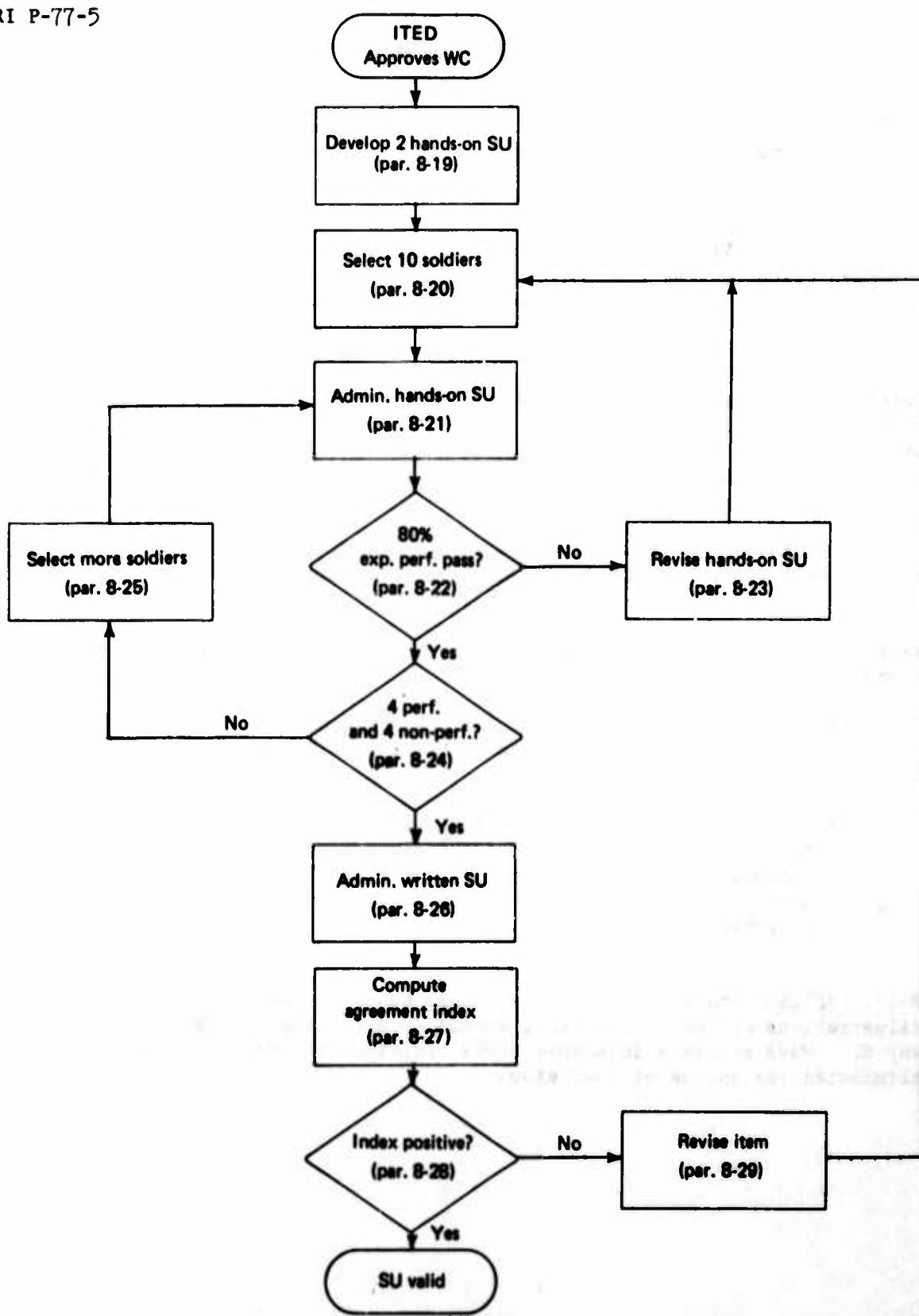


Figure 8-4. Validate Two Written SU Against Hands-On

Section III. VALIDATE TWO WRITTEN SU AGAINST HANDS-ON

8-18. INTRODUCTION. Each option in the validation plan requires you to select a group of soldiers who can do a task (performers) and a group who cannot do a task (non-performers). You will then compare the two groups on their written test scores for that task.

Options vary in the way you identify performers and non-performers. The best option for identifying performers and non-performers is to give a group of soldiers a hands-on test for the task. Unfortunately that is also the most expensive option.

It is not likely that any TDA will have the resources to validate all written SU against hands-on tests. But each TDA must validate at least two written SU against hands-on tests. TDA will thus build up a reservoir of written SU validated against the strongest possible criterion.

8-19. DEVELOP TWO HANDS-ON SU. Choose two tasks that you think will be in the SQT in the future. Develop a hands-on test for each task. Follow the procedure in Chapter 5.

8-20. SELECT TEN SOLDIERS. Select 10 soldiers in the MOS the SQT addresses. Assure a range of ability among the soldiers by getting five soldiers you expect to be performers (high rank or AIT instructors) and five you expect to be non-performers (low rank or AIT students).

If you cannot get 10 soldiers, coordinate with ITED for permission to reduce the size of the sample.

8-21. ADMINISTER HANDS-ON TEST. Test each soldier you select on the hands-on SU. Soldiers who receive a GO are performers; soldiers who receive a NO-GO are non-performers.

8-22. DO 80% OF EXPECTED PERFORMERS PASS? If the people you expect to be able to do the task fail the hands-on test, something may be wrong with the hands-on test. Discuss the test with expected performers to find out why they failed. If expected performers insist that the test is wrong, revise the test.

8-23. REVISE HANDS-ON SU.

a. If expected performers do not pass the SU and insist that they really are performers, two things might be wrong:

- The performance measures may be too imprecise.
- The test may be based on incorrect doctrine.

On the other hand, the test may be pointing up an error so common that even expected performers commit it.

b. If you find that the doctrine for the test is wrong, check the written SU for that task. If it is also based on incorrect doctrine, revise the written SU.

8-24. DOES SAMPLE INCLUDE AT LEAST FOUR PERFORMERS AND FOUR NON-PERFORMERS? The analysis for this validation requires that each group (performer and non-performer) be represented in the sample. At least four of each and a total of ten will meet the requirements. If more are needed in either group, select more soldiers.

8-25. SELECT MORE SOLDIERS. If either group is under-represented, select more soldiers in the MOS and administer the hands-on SU to them.

8-26. ADMINISTER WRITTEN SU. Have the 10 soldiers in the validation sample answer the questions in the written SU. Make conditions as close as possible to conditions for operational administration. Use the same illustrations and instructions. Allow soldiers as much time as needed to answer each question and to review their answers.

8-27. COMPUTE AGREEMENT INDEX.

a. First, score the written SU. Apply the scoring rules you developed for the SU.

b. Second, prepare a 2x2 table for each item. The table should contain the following information:

## HANDS-ON TEST

WRITTEN ITEM	Pass	Fail
Pass	$a = \text{Number of soldiers who pass hands-on test and pass written item.}$	$b = \text{number of soldiers who fail hands-on test and pass written item.}$
Fail	$c = \text{Number of soldiers who pass hands-on test and fail written item.}$	$d = \text{number of soldiers who fail hands-on test and fail written item.}$

$$\text{Agreement Index} = (a \times d) - (b \times c)$$

c. Third, compute the Agreement Index [AI =  $(a \times d) - (b \times c)$ ] for each table.

d. For example, assume that four of six performers (GO on hands-on test) pass an item that only one of four non-performers (NO-GO on hands-on test) passes. The 2x2 table would look like this:

4	1
2	3

The equation for the table is:  $AI = (4 \times 3) - (2 \times 1)$   
 $AI = 12 - 2$   
 $AI = 10$

The agreement Index of 10 indicates an acceptable level of agreement between performance on the hands-on test and performance on the item. Any index greater than zero is acceptable.

8-28. IS INDEX POSITIVE? If performers do better on an item than non-performers do, the index will be higher than zero. If the index is zero or a negative number, there is probably something wrong with the item. Such items must be revised. If the index is positive, the item is valid.

8-29. REVISE ITEM.

a. If the people who developed the SU have followed development procedures, minor revisions will probably make the items acceptable. If the item is too hard (more performers fail than pass it), check the wording of the item, check the accuracy of the correct alternative, and consider adding illustrations to increase clarity.

If the item is too easy (non-performers pass it at least as often as performers), check for unwanted grammatical cues, check for cues in other items that may give away the item in question, and check the possibility that a likely wrong alternative has been left out of the item.

b. If the item does not have a superficial deficiency, you may need to make a major revision. First analyze responses of the performers to find which wrong alternative they usually select. Be sure that alternative is a real-world alternative and is wrong. Then check the task analysis data to be sure that the behavior the item measures is a key behavior.

If the behavior is not key, you may be able to coordinate with ITED to delete it from the SU. If after careful study you can find no fault with the item and the behavior is key to successful performance of the task, accept the item as is.

c. Repeat the validation process for each item you revise.

Section IV. SELECT VALIDATION OPTION FOR WC

8-30. INTRODUCTION. Six options are available for validating the WC. The option you choose will depend on availability of incumbents in the MOS, supervisors of incumbents in the MOS, or experts in the MOS. The procedure for selecting a validation option is based on two assumptions.

The first assumption is that the self-rating option with 30 soldiers is the most practical option. We know more about it than the other two and it poses the fewest logistical problems. The second assumption is that most TDA will prefer to conduct the validation at their installation.

Therefore the procedure recommends considering all three local options before considering the TDY options. If you have TDY funds available, you may wish to alter priority of the options and, for example, select the TDY self-rating option before trying the local supervisor or local panel-of-experts options.

The intent of this procedure is to give TDA flexibility in validating the WC. You may combine the methods to validate the complete component. If even this flexibility is not enough, coordinate with ITED.

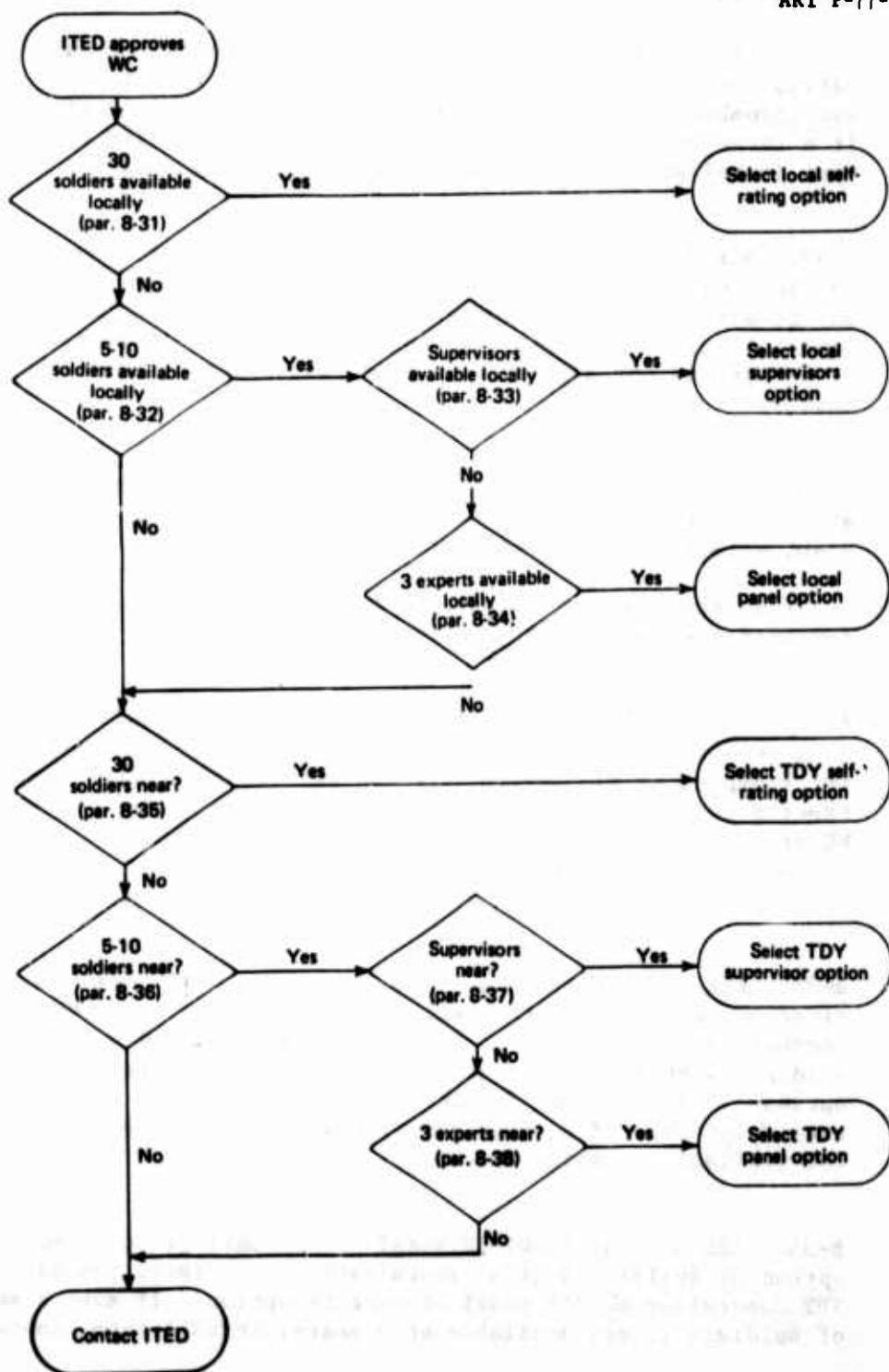


Figure 3-5. Select Validation Option for WC

8-31. ARE AT LEAST 30 SOLDIERS AVAILABLE LOCALLY? Soldiers for the validation should be incumbents in the MOS or a mixture of AIT students and incumbents. The goal is a spread of ability to perform the tasks. If a large group of soldiers with a spread of ability is available at your installation, select the local self-rating option.

8-32. ARE 5 TO 10 SOLDIERS AVAILABLE LOCALLY? If a small group of soldiers with a spread of ability is available at your installation, you may be able to select either the local-supervisor option or the local panel-of-experts option. If even a small group of soldiers with a range of ability is not available at your installation, consider the TDY options.

8-33. ARE SUPERVISORS AVAILABLE LOCALLY? If the small group of soldiers at your installation has at least one immediate supervisor you can interview, select the local supervisor option.

If no supervisor is available for the small group of soldiers, consider the local expert option.

8-34. ARE THREE EXPERTS AVAILABLE LOCALLY? For this decision, an expert is a person generally acknowledged to be a high-level performer on a task, for example, an AIT instructor. Do not consider other test developers experts for this decision. If at least three experts on each task in the WC are available at your installation, select the local panel of experts option. If not enough experts can form a panel for each SU, consider TDY options.

8-35. ARE 30 SOLDIERS NEAR? In this case "near" means an installation close enough to visit that will allow access to soldiers. If you visit another installation to conduct the validation and a large group of soldiers with a range of ability is available, select the TDY self-rating option. The procedure for this option is the same as for the local self-rating option. If a large group of soldiers with a range of ability is not available, consider other TDY options.

8-36. ARE 5 TO 10 SOLDIERS NEAR? If a small group of soldiers with a spread of ability is at an installation you visit, you may select the TDY supervisor or TDY panel of experts option. If even a small group of soldiers is not available at a nearby installation, contact ITED.

8-37. ARE SUPERVISORS NEAR? If the small group of soldiers at the nearby installation has at least one immediate supervisor you can interview, select the TDY supervisor option. The procedure for this option is the same as for the local supervisor option described earlier. If there is no available supervisor for the small group of soldiers, consider the TDY expert option.

8-38. ARE THREE EXPERTS NEAR? If at least three high level performers on each task in the WC are available at a nearby installation, select the TDY expert option. The procedure for this option is the same as for the local expert option described earlier. If sufficient experts are not available, contact ITED.

#### Section V. VALIDATE WRITTEN SU - SELF RATINGS

##### 8-39. SELECT 30 SOLDIERS.

a. Soldiers for this option may be incumbents in the MOS, or AIT students or a combination of incumbents and AIT students. Incumbents in the MOS do not have to be at the skill level for the SQT, in fact a range of skill levels is desirable since that helps assure a spread of ability to perform the tasks.

b. Have each candidate for the sample rate his or her own ability to perform the tasks covered by the WC. When you collect the self-ratings, emphasize the importance of accuracy. Be sure expected experts know there is no benefit to them or the SQT if they overrate their ability. Give each soldier a detailed description of the task that specifies the task conditions and standard. Then, ask each soldier one of the following questions. The question to ask depends on whether the task is normally performed under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

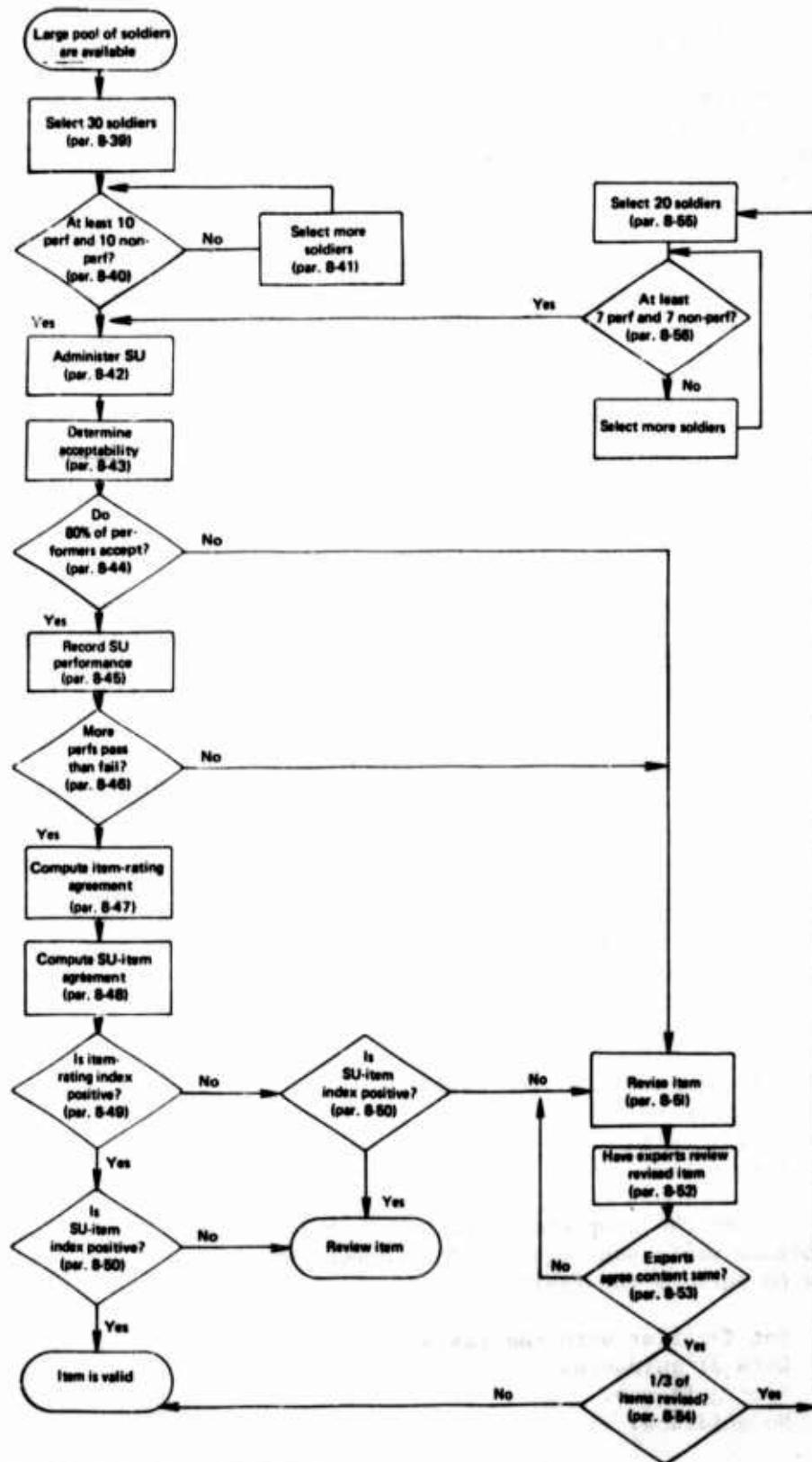


Figure 8-6. Validate Written SU - Self Ratings

c. The self-ratings identify groups of performers and non-performers for each SU. Soldiers who rate themselves low (0 or 1) are in the non-performer group; soldiers who rate themselves high (2 or 3) are in the performer group. A soldier may be in the performer group for one task and in the non-performer group for another task.

**8-40. DOES SAMPLE HAVE AT LEAST 10 PERFORMERS AND 10 NON-PERFORMERS?** One requirement for a valid item is that a greater proportion of performers than non-performers pass the item. To tell if that happens, you must be sure that the sample of soldiers includes performers and non-performers. Also assure that the size of each group is large enough to reduce impact of incorrect self-ratings. At least 10 per group and a total of 30 will enable you to be confident in the computed Agreement Index.

**8-41. SELECT MORE SOLDIERS.** If you have less than 10 performers or non-performers, add soldiers and collect self-ratings until you have at least 10 soldiers per group and a total of 30.

**8-42. ADMINISTER THE SCORABLE UNIT.** Have soldiers in the sample answer the questions in the SU. Make conditions similar as possible to field conditions. Use illustrations and instructions that will be used when you test for record. Allow soldiers as much time as they need to answer each question and to review their answers.

**8-43. DETERMINE ACCEPTABILITY.** After soldiers finish the SU, have them fill out Question Sheet 2 (Figure 8-7). The questions concern completeness and fairness of the SU. Although acceptability to non-performers is not a criterion for a valid item, performers and non-performers should complete the Question Sheet for two reasons:

- Feedback from non-performers will be helpful if you need to revise an item.
- Soldiers in the sample will probably cooperate better if you treat all soldiers alike regardless of their self-ratings.

**8-44. DO 80% OF PERFORMERS ACCEPT EACH ITEM?** The first requirement for an SQT item is that soldiers who say they can do a task think the item is a fair way to find whether a person can do an important part of the task. If more than 20% of the performers say a specific item is irrelevant or unfair, revise the item.

NAME \_\_\_\_\_

EVALUATION OF WRITTEN SCORABLE UNITS

Each item should be based on a behavior (decision or action) essential to performing the task.

Yes      No

1. Does passing each item correspond to a behavior (decision or action) necessary for task performance? \_\_\_\_\_

If not, please specify which items should be revised:

2. Do the items for each Key Behavior fully measure that point? (Key Behaviors have special consequences to personnel or equipment or are frequent sources of failure) \_\_\_\_\_

If not, please specify which items should be:

- a. Added

Which one(s)?

- b. Revised

Which one(s), and how?

3. Based on this Scorable Unit, would you be confident that soldiers who answer items correctly are able to perform the task? \_\_\_\_\_

Figure 8-7. Question Sheet 2

8-45. RECORD SU PERFORMANCE.

- a. Clear your desk, sharpen four pencils, check your eraser and tell your secretary to hold all calls from anyone who does not wear a star. Now the fun begins.
- b. First, score the tests. Apply the scoring rules you developed for each item in the SU.
- c. Second, separate the answer sheets for the performers (self-rating of 2 or 3) from the answer sheets for the non-performers (self-rating of 0 or 1). Remember, a soldier may be a performer on some tasks and a non-performer on another; therefore check the self-ratings for each SU.
- d. Third, for each SU list the self-rating, performance per item, and total items passed for each performer. The Ordnance School has developed an Item Analysis form you can use to make the list (Figure 8-8). Enter a "1" for each item a soldier passes and a "0" for each item he fails. (You could, of course, use "P" and "F" but numbers are more scientific.)
- e. Fourth, record the total number of performers who pass each item and the number who fail each item.
- f. Fifth, list the self-rating, performance per item and total items passed for each non-performer.
- g. Sixth, record the number of non-performers who pass each item and the number who fail each item.

8-46. DO MORE PERFORMERS PASS THAN FAIL EACH ITEM? The second requirement for an SQT item is that more soldiers who say they can perform the task pass the item than fail it. If an item does not meet this requirement, the item probably requires knowledge or a skill not part of the job. Revise such items to eliminate artificial difficulty.

8-47. COMPUTE ITEM-RATING AGREEMENT.

- a. The third requirement for an SQT item is that performers do better on the item than non-performers do. You can find whether that happens by computing an Agreement Index between performance on the item (pass/fail) and self-ratings on the task (performer/non-performer).

SCORABLE UNIT NO. \_\_\_\_\_ TASK NO. \_\_\_\_\_

Ident No. Examinee	Self Rating	Item Number										TOTAL	
		1	2	3	4	5	6	7	8	9	10	Pass	Fail
2	3	0	1	0								1	2
3	3	1	1	1								3	0
4	3	1	0	1								2	1
9	3	1	1	0								2	1
11	3	1	0	0								1	2
14	3	1	1	1								3	0
16	3	0	0	0								0	3
19	3	0	1	1								2	1
20	3	1	0	1								2	1
1	2	1	0	1								2	1
7	2	1	0	1								2	1
12	2	1	1	1								3	0
13	2	0	1	0								1	2
25	2	1	1	1								3	0
30	2	1	1	1								3	0
Pass		11	9	10									
Fail		4	6	5									
5	1	1	0	1								2	1
8	1	0	0	1								1	2
10	1	1	0	0								1	2
21	1	1	1	1								3	0
23	1	0	0	0								0	3
27	1	0	0	1								1	2
28	1	1	0	1								2	1
29	1	0	0	0								0	3
6	0	0	0	1								1	2
12	0	0	1	1								2	1
16	0	0	0	1								1	2
17	0	1	0	0								1	2
22	0	0	0	1								1	2
24	0	0	0	1								1	2
26	0	0	0	1								1	2
Pass		5	2	11									
Fail		10	13	4									

Figure 8-8. Item Analysis Form

b. The first step in computing an item-rating Agreement Index is to prepare a 2x2 table for the task. The table should contain the following information from the summary of performance:

## SELF-RATING

ITEM	Performer	Non-Performer
Pass	a = Number of soldiers who rate themselves as performers and pass item	b = Number of soldiers who rate themselves as non-performers and pass item
Fail	c = Number of soldiers who rate themselves as performers and fail item	d = Number of soldiers who rate themselves as non-performers and fail item

$$\text{Agreement Index} = (a \times d) - (b \times c)$$

c. The second step is to apply the formula:  $\text{Agreement Index} = (a \times d) - (b \times c)$ .

d. For example, consider item 1 from Figure 8-8. Eleven performers passed the item and four failed it; five non-performers passed the item and ten failed it. The 2x2 table is:

11	5
4	10

The equation for the table is:  $AI = (11 \times 10) - (4 \times 5)$   
 $AI = 110 - 20$   
 $AI = 90$

The Agreement Index of 90 indicates an acceptable level of agreement between performance and self-ratings; in fact, any index greater than zero is acceptable.

## 8-48. COMPUTE ITEM-SU AGREEMENT.

a. The fourth requirement for an SQT item is that soldiers who pass the item also pass the other items in the SU. The way to check whether that happens is to compute the Agreement Index between the item and the SU. The procedure to compute the Agreement Index depends on whether the SU contains up to four items or more than four items.

b. If the SU has four or less items, compute Agreement Indexes for all possible pairs of items in the SU. A two-item SU would have one pair; a three-item SU would have three pairs; and a four-item SU would have six pairs. Unless you really like to do things like this, do not compute agreement for items already marked for revision. The 2x2 table for calculating agreement for items 1 and 2 would contain the following information:

		ITEM 1	
		Pass	Fail
		a = Number who Pass Item 1 and Pass Item 2	b = Number who Fail Item 1 and Pass Item 2
ITEM 2	Pass	c = Number who Pass Item 1 and Fail Item 2	d = Number who Fail Item 1 and Fail Item 2
	Fail		

$AI = (a \times d) - (b \times c)$

Collect this information by referring to the summary of test performance and counting the number of times each of the four events occurs. The distinction between performers and non-performers does not apply for this analysis.

c. After preparing the 2x2 tables, apply the formula: Agreement Index =  $(a \times d) - (b \times c)$ .

d. For example, the 2x2 tables for the items in Figure 8-8 are shown in Figure 8-9. The 2x2 tables are prepared on another form developed at the Ordnance School. Again the minimum requirement is that the Agreement Index be greater than zero. All pairs in this sample SU meet that requirement.

e. If the SU has more than four items, compute Agreement Indexes between each item and overall performance on the SU. This computation requires four steps:

- Delete unacceptable items from total score. By this point, you may have already found an item that must be revised because 80% of the performers do not accept it, or because more performers fail than pass it. Including such items in this analysis will only distort the Agreement Indexes.
- Prepare a table indicating number of soldiers (performers and non-performers combined) who pass all items, all but one item, all but two, and so on down to those who pass zero items. The table for Figure 8-10 would look like this:

No. Items passed	6	5	4	3	2	1	0
No. of soldiers	2	5	8	8	2	4	1

(Note: Results of item 4 are dropped because too many performers failed it - remember the earlier requirement that more performers must pass an item than fail it.)

- Divide the sample into two groups as close as possible to the same size. In this example, the groups split evenly between 4 correct and 3 correct: the 15 soldiers who answered four or more questions right are the high group; the 15 soldiers who answered 3 or less questions right are the low group.
- Compute the Agreement Index between scoring group (high or low) and performance on each item (pass or fail). Prepare a 2x2 table for each item in this way:

#### SCORING GROUP

		High	Low
ITEM	Pass	a = Number of soldiers in high group who pass item	b = Number of soldiers in low group who pass item
	Fail	c = Number of soldiers in high group who fail item	d = Number of soldiers in low group who fail item

ARI P-77-5

SU No. \_\_\_\_\_ SQT \_\_\_\_\_ Task No. \_\_\_\_\_

ITEM 2			ITEM 3		ITEM 4	
	P	F	P	F	P	F
ITEM 1	P	7	9	12	4	
	F	4	10	9	5	
ITEM 2	P			8	3	
	F			13	6	
ITEM 3	P					
	F					

ITEM            PAIRS            AGREEMENT INDEX

ITEM 1 - ITEM 2  $(7 \times 10) - (9 \times 4) = 70 - 36 = 34$

ITEM 1 - ITEM 3  $(12 \times 5) - (4 \times 9) = 60 - 36 = 24$

ITEM 1 - ITEM 4

ITEM 2 - ITEM 3  $(8 \times 6) - (3 \times 13) = 48 - 39 = 9$

ITEM 2 - ITEM 4

ITEM 3 - ITEM 4

Figure 8-9. Scorable Unit Analysis Sheet - Four or Less items

SCORABLE UNIT NO. \_\_\_\_\_ TASK NO. \_\_\_\_\_

Ident No. Examinee	Self Rating	Item Number										TOTAL			
		1	2	3	4	5	6	7	8	9	10	Revised Pass	Pass	Fail	
2	3	1	1	1	1	1	0	1				5	6	1	
3	3	1	1	1	0	1	1	1				6	6	1	
4	3	1	1	1	0	1	1	1				6	6	1	
9	3	1	1	1	0	1	1	1				6	6	1	
11	3	1	1	1	0	0	1	1				5	5	2	
14	3	1	0	1	0	1	1	1				5	5	2	
18	3	1	0	1	0	1	1	1				5	5	2	
19	3	0	1	1	1	1	1	0				4	5	2	
20	3	1	1	0	0	0	1	1				4	4	3	
1	2	0	1	1	0	1	0	1				4	4	3	
7	2	1	0	1	0	0	1	1				4	4	3	
12	2	0	0	1	1	0	1	1				3	4	3	
13	2	0	1	1	1	0	1	0				3	4	3	
25	2	1	0	1	0	0	1	1				4	4	3	
30	2	1	0	1	0	1	1	0				4	4	3	
		Pass	11	9	14	4	9	13	12						
		Fail	4	6	1	11	6	2	3						
		5	1	1	0	1	0	0	0	1		3	3	4	
		8	1	1	0	1	0	0	1	0		3	3	4	
		10	1	1	0	1	0	0	1	0		3	3	4	
		21	1	0	0	1	0	1	1	0		3	3	4	
		23	1	1	0	1	0	0	1	1		4	4	3	
		27	1	1	0	1	1	0	0	0		2	3	2	
		28	1	0	0	0	1	0	0	1		1	2	0	
		29	1	0	0	1	0	0	1	0		2	2	0	
		6	0	1	0	1	0	0	0	0		2	2	0	
		12	0	0	0	1	0	0	0	0		1	1	0	
		16	0	0	0	0	0	0	0	1		1	1	0	
		17	0	0	0	1	0	0	0	0		1	1	0	
		22	0	0	0	0	0	0	0	0		0	0	0	
		24	0	0	0	1	0	0	1	1		3	3	0	
		26	0	1	1	1	0	0	0	0		3	3	0	
		Pass	7	1	12	2	1	6	5						
		Fail	8	14	3	13	14	9	10						

Figure 8-10. Scorable Unit Form

8-49. IS ITEM-RATING INDEX POSITIVE? If soldiers who rate themselves as performers on a task do better on an item than non-performers, the item-rating index will be greater than zero. If the index is zero or a negative number, probably something is wrong with the item. Such items must be reviewed or revised.

8-50. IS ITEM-SU INDEX POSITIVE? If high-scoring soldiers (high on overall SU) do better on an item than low scoring soldiers (low on overall SU), the SU-item index will be greater than zero. If the index is zero or a negative number, there may be something wrong with the item. If the item-rating index is not positive but the SU-item index is positive, review the item for possible revision. If neither index is positive, revise the item. If both indexes are positive and the other requirements are met, the item is valid.

8-51. REVISE ITEM.

a. If the people who developed the SU have followed development procedures, minor revisions will probably make the items acceptable. If the item is too hard (more performers fail than pass it), check the wording of the item, check accuracy of the correct alternative and consider adding illustrations to increase clarity.

If the item is too easy (non-performers pass it at least as often as performers), check for unwanted grammatical cues, check for cues in other items that may give away the item in question, and check possibility that a likely wrong alternative has been left out of the item.

b. If the item does not have a superficial deficiency, you may need to make a major revision. First analyze responses of the performers to find which wrong alternative they usually select. Be sure that alternative is a real-world alternative and is wrong.

Then check the task analysis data to be sure the behavior the item measures is a key behavior. If the behavior is not key, you may be able to coordinate with ITED to delete it from the SU.

If, after careful study, you can find no fault with the item, and the behavior is key to successful performance of the task, accept the item as is. When you decide whether to validate the revised SU, treat such an item as though it had been revised.

8-52. HAVE EXPERTS REVIEW REVISED ITEMS. Have at least three acknowledged experts compare the revised item with the original item. Ask them whether both items have the same content. Also have them answer the original and revised items.

8-53. DO EXPERTS AGREE THE CONTENT IS THE SAME? By this stage of the validation, content of the SU is established and cannot be changed without coordinating with ITED. If any expert disagrees with the answer or thinks the revision changed the content, revise the item again.

8-54. ARE 1/3 OF ITEMS REVISED? If fewer than 1/3 of the items in an SU are revised, no additional testing is required. If 1/3 or more of the items in an SU are revised, repeat the validation procedure for the complete SU.

8-55. SELECT 20 SOLDIERS. The only difference in the procedure for a revised SU is that your sample need contain only 20 soldiers with a range of ability. Do not include any soldiers who were in the original sample to validate the revised SU. Have soldiers rate their ability to perform the task.

8-56. AT LEAST SEVEN PERFORMERS AND SEVEN NON-PERFORMERS? The sample for validating revised SU, like the original sample, should include performers (self-ratings of 3 or 2) and non-performers (self-ratings of 0 or 1). If possible get 10 of each; have at least 7 of each. Perform the rest of the validation like you did originally.

#### Section VI. VALIDATE WRITTEN SU - SUPERVISOR RATING

8-57. INTRODUCTION. The supervisor rating validation option lets you validate a written component even though you have access to less than 30 soldiers.

This option selects the sample by combining a supervisor's rating of a soldier's observed or expected performance with the soldier's rating of his own ability to perform a task.

8-58. CONTACT SUPERVISOR WITH THE MOST PEOPLE IN THE MOS. Since you will need several soldiers in the MOS, it is a good idea to go to the supervisor responsible for the most soldiers in the MOS at the installation.

8-59. HAS SUPERVISOR OBSERVED SOLDIERS PERFORM THE TASK? Ask the supervisor if he has seen any of his soldiers, regardless of rank, try to perform the task covered by the SU. Show him the task analysis data for the task so he knows the scope of the task.

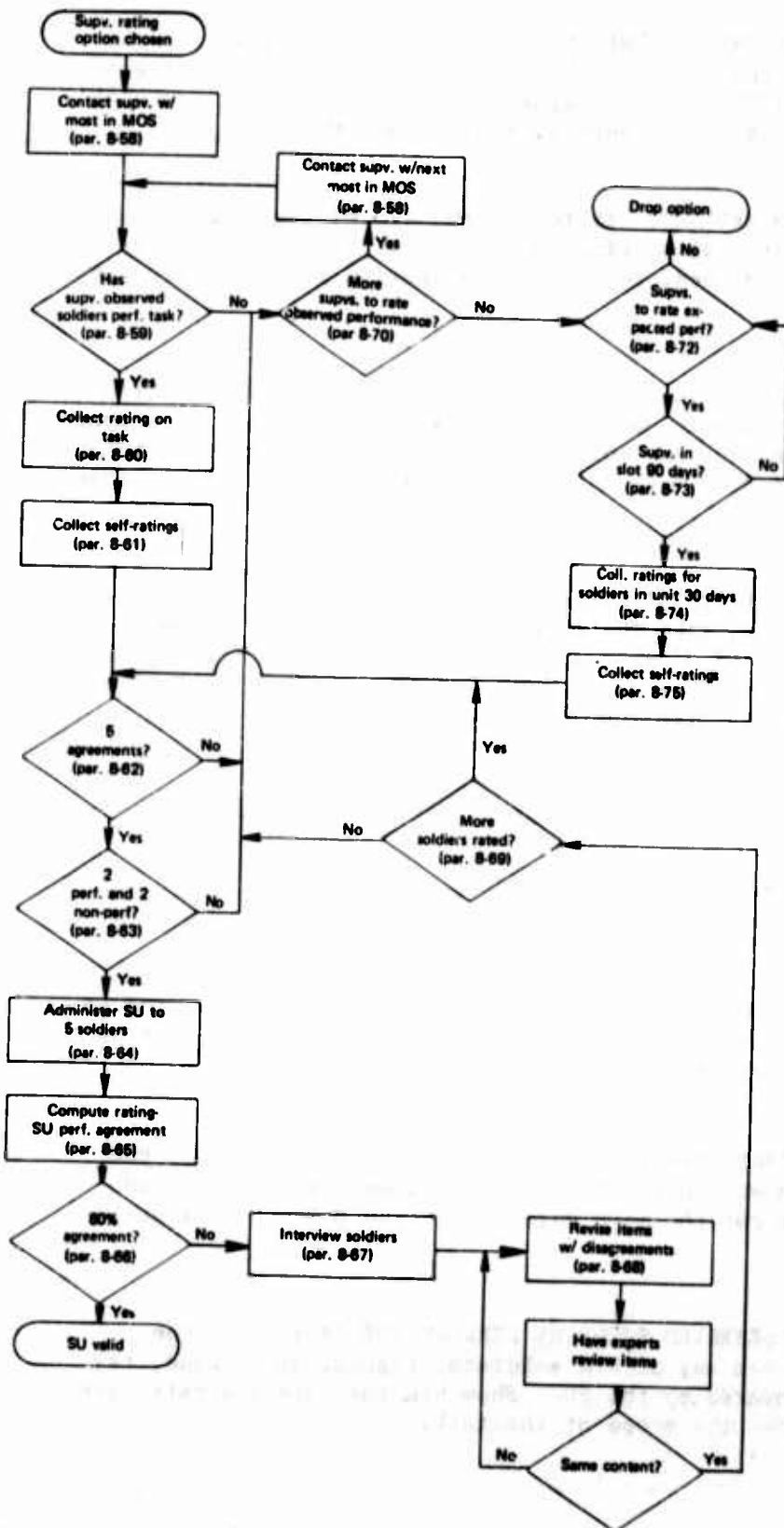


Figure 8-11. Validate Written Component-Supervisor Rating

If he has not observed at least five soldiers perform the task, eventually you will need to contact another supervisor, but go ahead and collect data for soldiers he has observed. If he has not observed any soldiers perform the task, tell him you may need to contact him later, but for now none of his soldiers are eligible to be in the sample.

8-60. COLLECT RATINGS ON TASK. Have the supervisor answer one of the following questions for each soldier he has seen perform the task. The question to ask depends on whether the task normally is performed under supervision.

(1) How well can he perform the task?

1. Not very well.
2. Fairly well.
3. Very well.

(2) To perform the complete task again, how much technical guidance would he require from someone who knows how to perform the task?

1. Lots of guidance.
2. Some guidance.
3. No guidance.

Soldiers rated "2" or "3" are performers. Soldiers rated "1" are non-performers.

8-61. COLLECT SELF-RATINGS. Collect self-ratings from each soldier the supervisor rates. When you collect the self-ratings, emphasize the importance of accurate self-ratings. Be sure expected experts know that there is no benefit to them or to the SQT if they overrate their ability. Give the soldier the task analysis data for the task and ask him one of the following questions. The question to ask depends on whether the task normally is performed under supervision.

(1) How well can you perform the task?

0. Not familiar with the task.
1. Not very well.
2. Fairly well.
3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

0. Not familiar with the task.
1. Lots of guidance.
2. Some guidance.
3. No guidance.

8-62. FIVE AGREEMENTS? Compare the supervisor's rating of each soldier with the soldier's self-ratings. Soldiers in the same group for both ratings are eligible to be in the validation sample. If less than five agreements, you will need to get more soldiers before you finish the validation, but complete the rest of the main strand (left side) of the flow-chart while the soldiers are available.

8-63. AT LEAST TWO PERFORMERS AND TWO NON-PERFORMERS? The five soldiers in the validation sample must have a range of ability. For this option, there must be at least two performers and two non-performers among the five. Otherwise you cannot tell whether an SU is too easy or too hard.

For example, if you have only non-performers in the sample and they all fail an SU, you still do not know whether performers could pass the SU. If you do not have a range of ability, eventually you will need to add to the sample, but test soldiers you need while they are available.

Do not test more soldiers in one group than you need. Save the "extra" soldiers in one group than you need. Save the "extra" soldiers in case the SU is revised.

8-64. ADMINISTER SU TO FIVE SOLDIERS.

a. Although eventually you must have five soldiers take each SU, there are two advantages to testing soldiers as they are available even though there are less than five of them:

- You do not have to arrange a time for soldiers from different units to get to a test site.
- You can be alerted to faults in the SU that will require revision. You can then revise the SU and conserve remaining soldiers for validating the revision.

b. When you administer the SU, make conditions as similar as possible to the field conditions. Use illustrations and instructions the same as when you test for record. Allow soldiers as much time as they need to answer each question and to review their answers.

8-65. COMPUTE RATING-SU AGREEMENT.

a. First, score the tests. Apply the scoring rules developed for each item in the SU.

b. Next prepare a table that shows each soldier's rating (performer or non-performer) and his performance on the SU (pass or fail). Here is what a table might look like:

Soldier	Rating	SU Score
1	1*	1+
2	0**	1
3	0	0++
4	1	1
5	0	0

\*1 = performer

+1 = pass

\*\*0 = non-performer

++0 = fail

c. Finally figure the percentage of agreement between ratings and SU scores. First count the number of agreements. An agreement is when a performer passes the SU or a non-performer fails the SU. In the example, two performers passed the SU and two non-performers failed it; therefore the example has four agreements. Then divide the number of agreements by the number of soldiers in the sample. The percentage of agreement in the example is 80 ( $4 \div 5$ ).

8-66. IS THE AGREEMENT AT LEAST 80%? A minimum agreement of 80 percent is required for the SU to be acceptable. With five to nine soldiers in the sample, an SU is allowed only one disagreement. If the agreement is 80 percent or higher, the SU is valid. If the agreement is less than 80 percent, you must revise the SU.

8-67. INTERVIEW SOLDIERS. If the SU does not validate (less than 80 percent agreement), interview each soldier one at a time to identify faults in the SU. Ask these questions:

- (1) Do you think soldiers who answer these items correctly are able to perform the task?
- (2) Does each question relate to an important part of doing the task?
- (3) Are any important parts of doing the task left out?

Then try to find why a performer failed an item or a non-performer passed the item. That information will help you revise the SU.

8-68. REVISE ITEMS WITH DISAGREEMENTS.

If performers do not pass an item, it may be too hard. Check wording of the item, check accuracy of the correct alternative, consider adding illustrations to increase clarity and check the wrong alternatives performers choose to be sure those alternatives are real-world choices and are really wrong.

If non-performers pass an SU the items may be too easy. Check for unwanted grammatical cues, check for cues in other items that may give away the item in question, and check possibility that a likely wrong alternative has been left out of the item.

8-69. ARE MORE SOLDIERS RATED? This is the first question to answer when you validate a revised SU. It is possible that you may have some soldiers left from the first rating whose self-rating agree with their supervisor's, but who were not needed for the first validation. Do not include soldiers from the first validation in the sample for the revised SU. The sample for a revised SU must meet the original requirements, that is, at least two performers and two non-performers among a group of five soldiers.

8-70. ARE MORE SUPERVISORS AVAILABLE TO RATE OBSERVED PERFORMANCE? If you cannot meet requirements for the original validation sample or for validating the revised SU based on observed performance from the first supervisor you contact, find if there are other supervisors of soldiers in the MOS at the installation. If there are, contact the supervisor with the next most in the MOS. If you have exhausted the possibilities of observed performance, collect ratings of expected performance.

8-71. CONTACT SUPERVISOR WITH NEXT MOST IN MOS. If you have not yet contacted all the supervisors, contact the remaining supervisor who is responsible for the most soldiers in the MOS. Continue contacting supervisors until you meet the sample requirements or run out of people who can rate on observations of performance.

8-72. ARE THERE SUPERVISORS TO RATE EXPECTED PERFORMANCE? A less powerful but still acceptable basis for supervisor's ratings is expected performance. Contact the supervisors of soldiers in the MOS who could not be rated on observed performance. Usually you will simple re-contact each supervisor.

8-73. HAS SUPERVISOR BEEN IN THE SLOT FOR 90 DAYS? The first requirement for expected performance ratings is that the rater be an experienced supervisor. The minimum level of experience is 90 days. If the supervisor has not been in his slot for 90 days, contact another supervisor.

## 8-74. COLLECT RATING FOR EACH SOLDIER WHO HAS BEEN IN THE UNIT 30 DAYS.

a. The second requirement for expected performance ratings is that the supervisor is familiar with the overall ability of soldiers he rates. If the soldier has been under the rater's supervision for 30 days, this requirement is met.

b. Give the supervisor the task analysis data for the task and have him answer this question for each soldier eligible to be in the validation sample:

(1) How confident are you that he can do this task?

- 0. Sure he cannot do it.
- 1. Doubt he can do it.
- 2. Think he can do it.
- 3. Sure he can do it.

Soldiers rated 0 or 1 are non-performers. Soldiers rated 2 or 3 are performers.

## 8-75. COLLECT SELF-RATINGS.

a. Collect self-ratings from each soldier the supervisor rates. Give the soldier the task analysis data for the task and ask him one of the following questions. The question to ask depends on whether the task normally is performed under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well.
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

Soldiers who choose "0" or "1" are non-performers; soldiers who choose "2" or "3" are performers.

b. The remaining procedure for validating an SU using supervisor's ratings of expected performance is the same as for ratings of observed performance.

ARI P-77-5

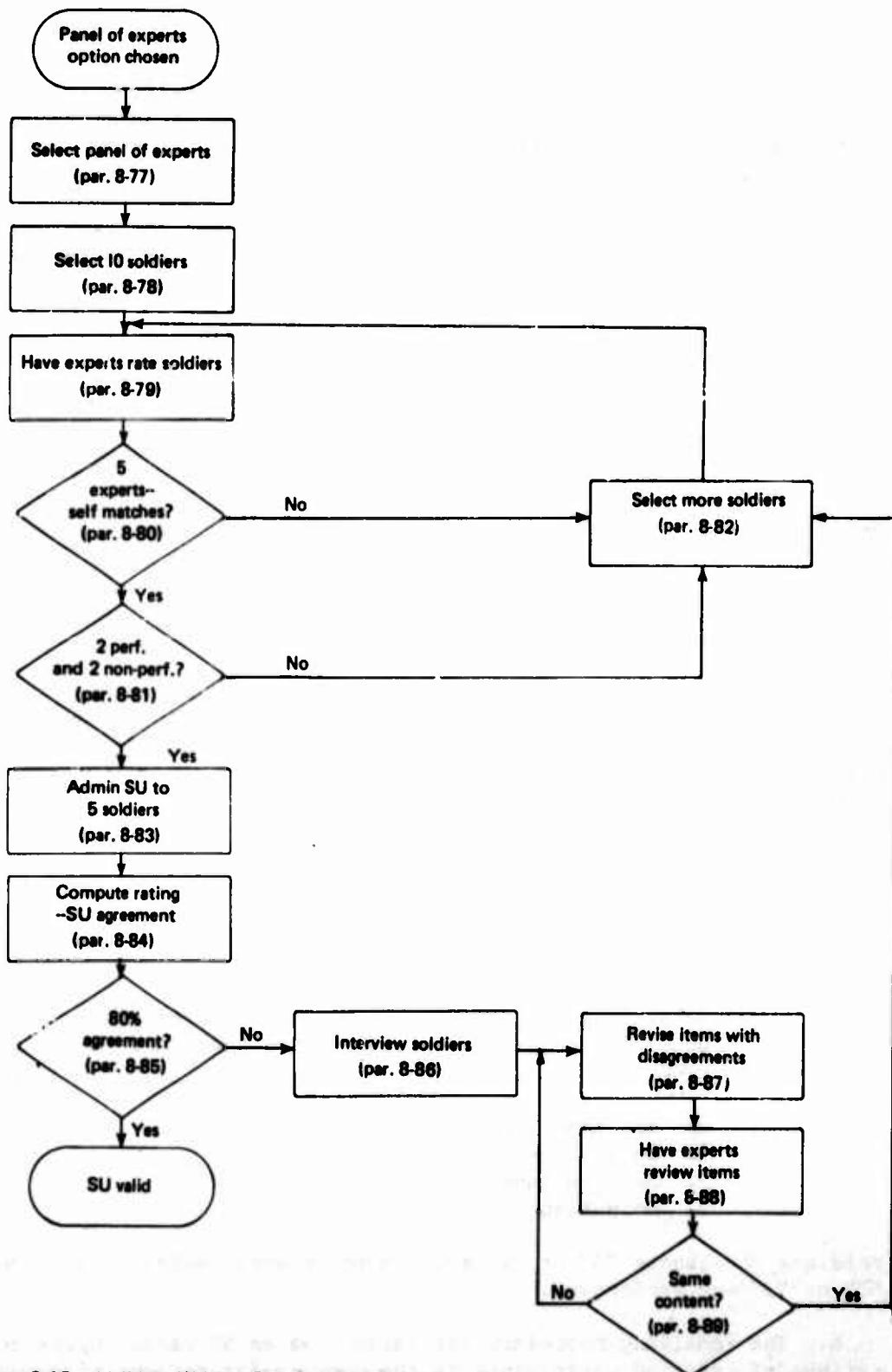


Figure 8-12. Validate Written SU - Panel of Experts

**Section VII. VALIDATE WRITTEN SU - PANEL-OF-EXPERTS**

8-76. **INTRODUCTION.** This option lets you validate an SU with fewer soldiers than the self-rating option even though supervisors of the soldiers are not available. With this option, you select the validation sample by combining the evaluation by a panel of experts and self-ratings made by 5 or 10 soldiers.

**8-77. SELECT PANEL OF EXPERTS.**

a. The panel should include three people who know a lot about the task covered by the SU and can do the task well. Two sources of experts are senior NCO's and instructors in the MOS. If you doubt the qualifications of a potential expert, interview his peers and supervisor. If they say the expected expert can do the specific task very well, consider including him on the panel. Otherwise do not include him on the panel. Test developers should not be on the panel.

b. After you locate several people you expect to be experts, give them the task analysis data for the SU and have them answer one of these questions. When you collect the self-ratings, emphasize the importance of accurate self-ratings. Be sure expected experts know that there is no benefit to them or the SQT if they overrate their ability. The question to ask depends on whether the task normally is performed under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well.
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

People who choose number 3, "Very well" or "No guidance" are eligible to serve on the panel.

8-78. SELECT SOLDIERS.

a. Soldiers for this option may be incumbents in the MOS, or AIT students, or a combination of incumbents and AIT students. Incumbents in the MOS do not have to be at the skill level for the SQT, in fact, a range of skill levels is desirable since that helps assure a spread of ability to perform the tasks.

Although the final sample requires only 5 soldiers, try to have at least 10 available because it is rare when everyone eligible meets requirements for the sample.

b. Give the soldiers the task analysis data for the SU and have them answer one of the questions. When you collect self-ratings, emphasize the importance of accuracy. Be sure expected experts know there is no benefit to them or the SQT if they overrate their ability. The question to ask depends on whether the task normally is performed under supervision.

(1) How well can you perform the task?

- 0. Not familiar with the task.
- 1. Not very well.
- 2. Fairly well.
- 3. Very well.

(2) To perform the complete task, how much technical guidance would you require from someone who knows how to perform the task?

- 0. Not familiar with the task.
- 1. Lots of guidance.
- 2. Some guidance.
- 3. No guidance.

Soldiers who rate themselves low (0 or 1) are in the non-performer group; soldiers who rate themselves high (2 or 3) are in the performer group. A soldier may be in the performer group for one task and in the non-performer group for another task.

8-79. HAVE EXPERTS RATE SOLDIERS.

a. The panel will interview each soldier to determine whether the soldier is a performer or non-performer on each task. The panel will develop their own interview procedures and questions based on the task analysis data. The panel is free to ask any questions they desire about task performance. The panel can ask for locations, identifications, procedures, special techniques and tools. The members should probe deeply by asking follow-up questions based on a soldier's responses.

b. Before the panel interviews anyone for record, have them rehearse their procedure. If you are short of soldiers, have members of the TDA role-play as potential subjects. Monitor the rehearsal to be sure that the panel does not appear threatening to soldiers it interviews.

c. Have the panel interview one soldier at a time. Each expert on the panel will classify each soldier as either a performer or a non-performer. In case of disagreement among the three experts, the rating that two experts agree on will be accepted.

8-80. AT LEAST FIVE EXPERT - SELF-RATING AGREEMENTS? Compare the experts' rating of each soldier with the soldier's self-rating. Soldiers who are in the same group (performer or non-performer) for both ratings can be in the validation sample. If less than five agreements, get more soldiers.

8-81. AT LEAST TWO PERFORMERS AND TWO NONPERFORMERS? The validation sample must include a range of ability. Otherwise, you cannot tell whether an SU is too easy or too hard. For example, if you have only non-performers in the sample and they all fail the SU, you still do not know whether performers could pass the SU. The minimum requirement is two in each group.

8-82. SELECT MORE SOLDIERS. If you have less than two performers and two non-performers, add soldiers, collect self-ratings and have experts rate soldiers until you have at least two soldiers per group and a total of five.

8-83. ADMINISTER SU TO FIVE SOLDIERS. When you administer the SU, make the conditions as similar as possible to the field conditions. Use illustrations and instructions the same as when you test for record. Allow soldiers as much time as they need to answer each question and to review the answers.

8-84. COMPUTE RATING-SU AGREEMENT.

a. First, score the tests. Apply the scoring rules developed for each item in the SU.

b. Next prepare a table that shows each soldier's rating (performer or non-performer) and his performance on the SU (pass or fail). Here is what a table might look like:

Subject	Rating	SU Score
1	1*	1+
2	0**	1
3	0	0++
4	1	1
5	0	0

\*performer

+pass

\*\*nonperformer

++fail

c. Finally figure the percentage of agreement between ratings and SU scores. First count the number of agreements. An agreement is when a performer passes the SU or a non-performer fails the SU. With the example, two performers passed the SU and two non-performers failed it; therefore the example has four agreements. Then divide number of agreements by number of soldiers in the sample. For the example, the percentage of agreements is 80 ( $4 \div 5$ ).

8-85. IS THE AGREEMENT AT LEAST 80%? A minimum agreement of 80 percent is required for the SU to be acceptable. With five to nine soldiers in the sample, an SU is allowed only one disagreement. If the agreement is 80 percent or higher, the SU is valid. If the agreement is less than 80 percent, revise the SU.

8-86. INTERVIEW SOLDIERS. If the SU does not validate (less than 80 percent agreement), interview each soldier one at a time to identify faults in the SU. Ask these questions:

- Do you think soldiers who answer these items correctly are able to perform the task?
- Does each question relate to an important part of doing the task?
- Are any important parts of doing the task left out?

Try to find why a performer failed an item or a non-performer passed the item. That information will help when you revise the SU.

8-87. REVISE ITEMS WITH DISAGREEMENTS. If performers do not pass an item, it may be too hard. Check wording of the item, check accuracy of the correct alternative, consider adding illustrations to increase clarity and check the wrong alternatives performers choose to assure those alternatives are real-world choices and really wrong. If non-performers pass an SU, the items may be too easy. Check for unwanted grammatical cues, check for cues in other items that may give away the item in question and check possibility that likely wrong alternatives have been left out of an item.

8-88. HAVE EXPERTS REVIEW REVISED ITEMS. Have at least three acknowledged subject matter experts compare the revised item with the original item. Ask them whether both items have the same content. Also have them answer the original and revised items.

8-89. EXPERTS AGREE THE CONTENT IS THE SAME?

a. By this stage of the validation, content of the SU is established and cannot be changed without coordinating with ITED. If any expert disagrees with the answer or thinks the revision changed the content, revise the item again.

b. If experts agree the content is the same, validate the revised SU. Perform the same procedure described for the original validation.

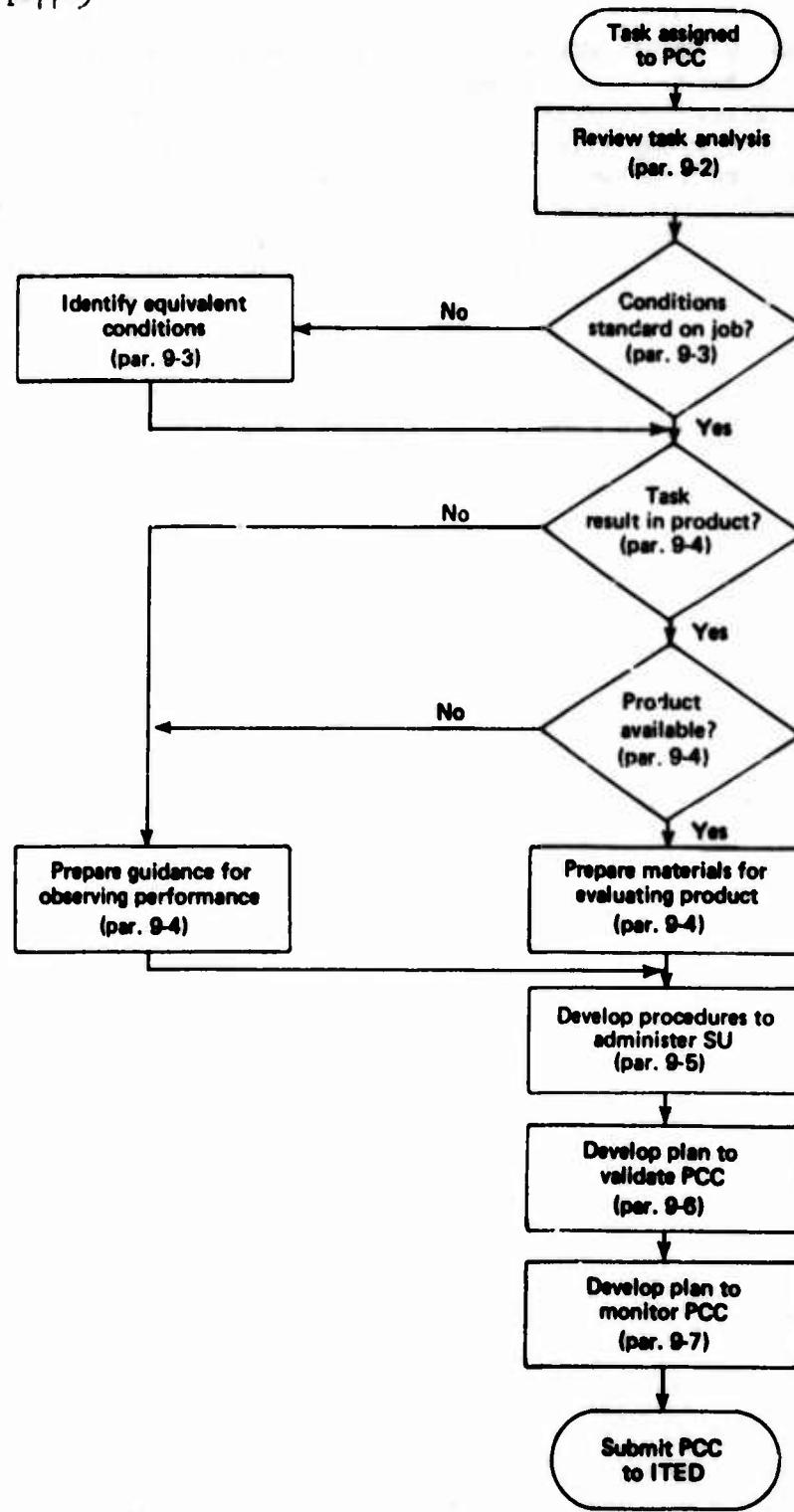


Figure 9-1. Develop Performance Certification (PCC) SU

## CHAPTER 9

### DEVELOP PERFORMANCE CERTIFICATION (PCC) SU

#### 9-1. INTRODUCTION.

a. The allocation of tasks to the Performance Certification Component (PCC) was discussed in chapter 4. Despite the serious weaknesses of the PCC as a means of measuring a soldier's proficiency, it appears the only way that certain highly skilled performances can be evaluated.

Performance Certification is the component where you should consider testing a task that

(a) cannot be validly tested in the WC of its skilled hands-on requirements, and

(b) cannot be tested in the HOC because of administrative constraints.

Sometimes a hands-on skill simply takes too long to demonstrate, or requires equipment, terrain or other conditions that prevent its testing in the HOC.

Cutting a roadbed with a bulldozer is a good example. To test this task in the HOC would probably take several hours and a sizeable expanse of terrain. Moreover, equipment and fuel requirements probably could not be justified merely for testing purposes.

b. The thinking behind the PCC is that such tasks can be "tested" on the job. By moving the test to the job site, we can have the soldiers' supervisor serve as scorer in evaluating the soldier's performance during the normal course of his job duties.

Some natural PCC already exist. Annual rifle qualification is one such example, where a soldier's task proficiency is evaluated in the field, but under conditions essentially the same as one would set up for a hands-on test. And this task is being included in the PCC for relevant MOS.

Most tasks that are candidates for the PCC, however, will not be structured already for testing in a field unit, so the test developer must prepare a scoresheet and carefully phrased instructions for the evaluator (supervisor) on how, when, where and under what conditions a soldier's task performance is to be scored.

c. Before discussing procedure for developing a PCC scorable unit, weaknesses inherent in the PCC should be noted carefully. These weaknesses are of two general types - both involving problems of standardization. The first pertains to difficulties in achieving standard conditions for task performance for all soldiers. In setting up a test, as was done in the HOC, we took great care in insuring that conditions were the same for each soldier going through the test station. In the PCC, where a soldier's task performance will be scored as part of his normal job duties, it will be nearly impossible to control the equality of conditions from man to man, location to location. A second weakness in the PCC is that the scorer (supervisor), who is beyond control of a Test Control Officer, in most cases will not be as well trained or as motivated to do a conscientious and unbiased job of scoring as a scorer in the HOC. The supervisor probably will have formed likes and dislikes among his subordinates. Just as serious is the possibility that a supervisor won't go to the trouble to score task performance by carefully controlled observation, but simply sit in his office and fill out the scoresheets on the basis of his preconceived notions of how the scores should come out.

d. These shortcomings of the PCC cannot be overcome entirely, which is why a task should be assigned to the PCC only as a last resort.

But, where a PCC is necessary, one should make every effort to (a) define the limits of variation in performance conditions that can be tolerated, and (b) make the supervisor's scoring requirements as simple, objective, and feasible as possible.

e. Keep in mind that the PCC, with its inherent problems in administration and support may not be tested in all locations. While every effort should be made to select items that are widely testable, there is no requirement that all soldiers take the PCC. Those soldiers that do not will receive an "N" score and will not be penalized for not having taken the PCC.

f. There is no requirement that the PCC be administered only once to a soldier. In fact, soldiers who fail an initial PCC should be allowed to be evaluated repeatedly until a passing score is obtained or it is evident that their score reflects their actual capability. This repetition will not be feasible in all cases but where it is, it should be encouraged.

9-2. REVIEW TASK ANALYSIS. Work begins with a careful review of the task analysis (see Chapter 3) just as it does in developing an SU for the HOC or the WC. In fact, the procedure for developing a scorable unit for the PCC is essentially the same as developing a SU for the Hands-On Component.

The major difference is that the test conditions and evaluator's perspective can vary more than is acceptable for the HOC. Therefore, when you develop a PCC scorable unit, you must prepare detailed guidance for evaluating the task under the various acceptable conditions.

#### 9-3. ARE CONDITIONS STANDARD ON THE JOB?

a. The first step is to specify acceptable conditions. Sometimes, essential conditions on the job will be the same. The mile run on the APEE will be essentially the same everywhere. But the PCC can still be valid even if the conditions vary on the job.

If conditions are not always the same, describe each aspect of the job condition that must be present before job performance is rated. Consider the amount of behavior, e.g., grade 100 yards of roadbed; the quality of the materials worked with, e.g., shale, sand, tundra; weather conditions (rain, snow, heat); and amount of supervision.

You may end up with several mixes of aspects of the job conditions. It is reasonable to require less behavior and allow more supervision when materials worked with and weather are more "difficult" than normal.

b. Review descriptions of conditions with a subject matter expert. The review should have two goals:

- (1) Be sure the conditions and standards are made as equivalent as possible.
- (2) Be sure you include all conditions under which competent incumbents could be expected to perform the task.

#### 9-4. DOES TASK RESULT IN A PRODUCT?

a. The second step in developing a PCC scorable unit is to decide how to evaluate a soldier's proficiency under each set of acceptable conditions. The method of evaluation depends first on whether the task results in a product. If it does, you must decide whether the commander will have or can find an existing representative product for each soldier who has performed the task during the preceding year. If representative products are available, the commander should evaluate the products. Prepare a scoresheet supplemented by a template, if possible, that will describe an acceptable product for each condition. The procedure for developing a product-scored SU for the HOC should be followed here.

b. If the task results in a product that may not be available for scoring, or if the task requires process measurement, prepare a score-sheet for the commander or his representative to observe the performance of the task.

If the procedure varies under some conditions, prepare a separate scoresheet for each set of conditions which causes the procedure to vary. In developing a process-scored SU for the PCC, the corresponding guidance for a process-scored HOC should be followed.

c. A sample process-scored SU with scoring instructions is shown at Appendix A.

**9-5. DEVELOP PROCEDURES TO ADMINISTER SU.** After you prepare guidance and materials for scoring an SU, you must develop procedures to administer the SU. These procedures should answer five questions:

- (1) How many trials will each examinee be allowed to demonstrate mastery?
- (2) If more than one trial is allowable, how much time should be between trials?
- (3) Who (specific duty position) will certify mastery, that is, score the test?
- (4) Who will fill out the mark sense form?
- (5) How will each soldier know he is being tested?

Submit the procedures for administering the SU to ITED at the same time you submit the PCC. These materials should accompany the SQT sample.

**9-6. DEVELOP PLAN TO VALIDATE PCC.**

a. Every SU in the SQT must be validated. This requirement applies as much to the PCC as to any other component. Some tasks may be in the PCC because there is not enough equipment to administer a hands-on SU in the unit, but there is enough equipment to test five soldiers on the SU. In such cases, follow the procedure in chapter 6.

b. If you cannot validate the SU by following the procedures for the HOC, you must develop another validation plan. At a minimum, your validation plan should call for submitting the PCC materials to at least three commanders and five representative incumbents. Interview the commanders and incumbents to answer the following questions:

- (1) Are task boundaries accurate?
- (2) Is task critical at MOS skill level?
- (3) Can soldiers be observed doing task? (Commanders only.)
- (4) Can scoring be standard in specific unit? (Commanders only.)

c. Submit the plan (not the results) for validating each SU in the component to ITED at the same time you submit the PCC. These materials should accompany the SQT sample.

#### 9-7. DEVELOP PLAN TO MONITOR PCC.

a. Because test conditions for the PCC are so variable, the TDA must propose a procedure for monitoring the PCC. This procedure should protect against two unlikely, but possible, abuses.

The first abuse is that a commander may "protect" the unit by certifying everyone in the unit as a GO. The second abuse is that a commander may evade the requirement entirely by marking everyone as Not Rated even though the task is routinely performed.

b. The recommended procedure for monitoring a PCC is to spot-check ability of soldiers in some of the units. The units checked should be those that appear to have an unreasonable number of "GO" or "Not Rated" scores plus a random sample of other units.

c. The plan for implementing the recommended procedure for monitoring the PCC should contain three parts. In the first part, specify the criteria for identifying units to be checked. For example, you may be confident that at least one-third of a unit's maintenance personnel will change a gun tube during the test period.

A higher "Not Rated" level will alert you to a likely abuse. In the second part, specify how you propose to check the units. Usually the check will be a small scale performance test. But also consider approaches such as interviewing incumbents or checking maintenance records.

In the third part of the plan tell who will do the check. Probably the most common proposal will be to send a team from the TDA. But you should also consider approaches such as working through the Maintenance Assistance Instruction Team at an installation.

d. Submit the plan for monitoring the PCC to ITED at the same time you submit the PCC. These materials should accompany the SQT sample.

ARI P-77-5

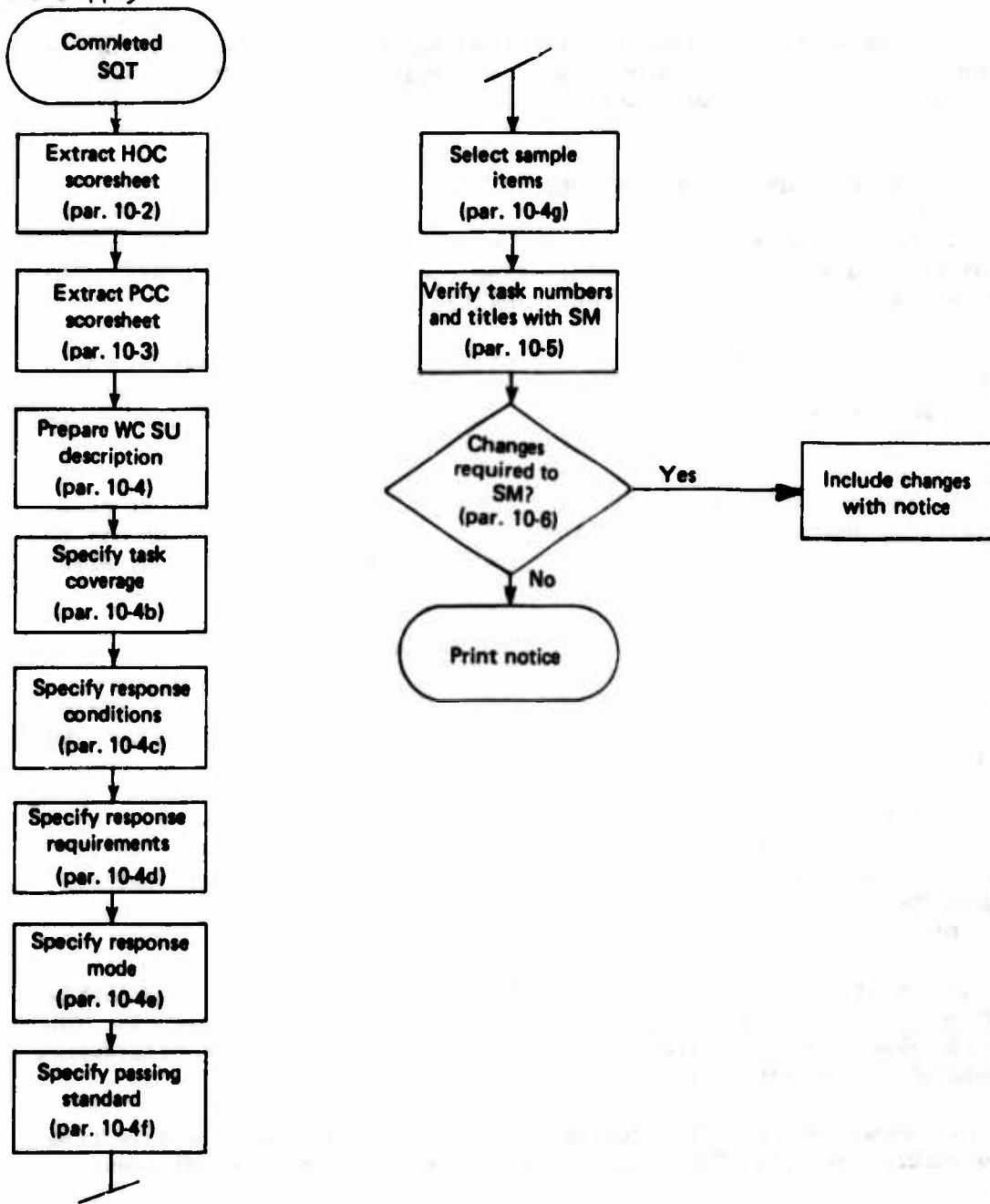


Figure 10-1. Prepare SQT Notice

## CHAPTER 10

## PREPARE SQT NOTICE

10-1. INTRODUCTION. The purpose of the SQT Notice is to tell the soldier precisely on what he will be tested. Unlike previous MOS tests, the soldier will be told exactly what he must know and be able to do to pass the test. For the HOC and PCC, this is accomplished by providing a copy of the scoresheet for each SU in the notice. Descriptions of written SUs are more difficult to prepare since we want the soldier to know what he will be asked to do without actually giving away the answers. How to prepare an entry for the Notice on a written SU is discussed in some detail. In addition, a description of the overall format for the Notice, a checklist for written SU entries, and a sample SQT Notice are given in Appendix B.

10-2. EXTRACT HOC SCORESHEET. The SQT Notice for the HOC will consist of an extract of the instructions as they are to be read to the examinee and the performance measures as they appear on the scoresheet. This information is extracted verbatim for the SQT Notice.

Be sure no special instructions to the scorer were included on the scoresheet that should not have been, e.g., the equipment fault to be induced in a trouble-shooting task. Such information should have been contained in instructions to the scorer, separately from the scoresheet.

Compromising information of this sort must be deleted from the Notice and, for that matter, from the scoresheet.

10-3. EXTRACT PCC SCORESHEET. The Notice entry for a PCC SU is prepared exactly the same as for the HOC. Preparation is essentially a matter of reproducing the appropriate scoresheet.

## 10-4. PREPARE WC SU DESCRIPTION.

a. Description. The entry in the Notice for a written SU consists of a short descriptive paragraph summarizing what the soldier will be asked to do, what he will be given to do it with, how he will be expected to respond, and what is required to pass.

b. Specify Task Coverage. It may be necessary to tell the soldier what aspect or part of the task is to be covered in the SU. In most cases, it will be evident to the soldier from the task title and the description of response conditions and requirements that you will provide.

c. Specify Response Conditions. State everything the soldier will be given in the SU: how the problem, situation or questions will be framed; media used (photos, tables, charts, etc.); and any tools or aids he will be using.

For example: "You will be given a map protractor, grid scales, and selected parts of maps..."; or, "You will be given a partially completed DA Form 2400 (Equipment Utilization Record) and a blank DA Form 2404 (Equipment Inspection/Maintenance Worksheet)."

d. Specify Response Requirements. Next, tell the soldier what specific task actions or products he will be performing or evaluating.

For example: "You must carry out the necessary calculations and conversions asked for..."; "You will be asked to identify where entries are made on DA Form 2400 and 2404..."; or, "You must match the correct shape and color with the words that complete the [NBC] marker."

e. Specify Response Mode. Next, if not already evident from the foregoing statement of response requirements, tell the soldier how he will be responding. In most cases, this will be simply a matter of saying, "You must select the correct...[action] [treatment] [picture] [sequence] [answer]...from those listed."

f. Specify Passing Standard. Finally, tell the soldier the number of items in the SU and the number he must pass to receive a "GO." For example, "This unit will consist of five questions. You must answer four questions correctly to score a "GO" for the unit."

g. Select Sample Items. Even a detailed description of a written SU may not tell the soldier exactly what he must know and be able to do to pass the SU. This is especially true of performance-based SU.

If the description leaves any doubt about the requirements to pass the SU, add a sample item to the description. If the item requires an illustration for the SQT, include the illustration in the description. Generally, consider including a representative item in the SQT Notice for every performance-based SU.

10-5. VERIFY TASK NUMBERS AND TITLES WITH SM. Check to be sure that the title and number of each task in the Notice are exactly the same as given in the SM.

10-6. ANY CHANGES TO MATERIAL IN SM? During SQT development you may have corrected a significant error in a task analysis, divided a large task into subtasks, or even discovered a critical task that somehow failed to appear in the SM. If any changes of this sort affect items that appear in the SQT, write them up and distribute them with the SQT Notice.

## APPENDIX A

HEADQUARTERS,  
DEPARTMENT OF THE ARMY  
WASHINGTON, DC, 20310

PERFORMANCE CERTIFICATION COMPONENT  
SQT 2, MOSC 16J10  
SCORING INSTRUCTIONS TO SUPERVISORS

1. A performance certification component is included in the SQT in order to determine whether soldiers have certain skills and abilities which are important but which cannot be tested. For example, in the SQT for Engineer Heavy Equipment Operator, it is not possible to determine from a written test whether a soldier has the ability to grade a roadbed. It is not practical to set up a hands-on test to measure this ability but soldiers who perform the task on the job can be observed by a supervisor and certified as able (or not able) to perform the task satisfactorily.
2. The supervisor who observes the soldier's performance on the tasks in the performance certification is responsible for scoring the soldier. The scoring must be done by a supervisor in the grade of E-6 or above. In order to certify the soldier's performance, performance of the task must have been observed within the 12-month period preceding the beginning of the quarter in which the SQT is scheduled. The supervisor must have clear knowledge of the soldier's performance. Soldiers performing the task during the preceding 12 months who can be scored by a supervisor will be scored in accordance with the standards on the Performance Certification Scoresheet. If the task was not performed during the required time period or if the supervisor does not have a clear knowledge of the soldier's performance, then the soldier is to be scored "N" (not observed). The soldier will not be penalized for a score of "N."
3. a. SQT 2, MOSC 16J10 consist of a written component and a performance certification component.

- b. The performance certification component consists of one event.

<u>TASK NUMBER</u> <u>SOLDIER'S MANUAL</u>	<u>PAGE NUMBER</u>	<u>TASK</u>
1. 441-16J-1039	1-5	ENERGIZE THE ELECTRONIC SEARCH CENTRAL

4. A scoresheet page must be completed for each soldier holding MOSC 16J10. Print the soldier's last name, first name, and middle initial. Enter the nine-digit SSN as it appears on the soldier's social security or identification card. Enter the TCO number which is to be provided by the TCO. The SQT number for soldiers holding MOSC 16J10 is 16J2476.

APPENDIX A (continued)

5. The information on the scoresheet is to be transcribed to an attached mark sense answer sheet which must be returned to the TCO not later than (date). The performance certification component scoresheet is to be given to the soldier.

6. In transcribing the identifying information to the mark sense answer sheet, print the soldier's name on the mark sense sheet exactly as it appears on the scoresheet and then under each letter blacken the space corresponding to that letter. Print the SSN on the mark sense form and then blacken the spaces corresponding to the numbers. Follow the same procedure in transcribing the TCO number and the test number to the mark sense form.

7. The test scores are to be transcribed from the scoresheet to the mark sense answer sheets as follows:

a. The scores for Scorable Unit 1 are to be entered on the Mark Sense answer sheet at Unit 1. If the event was not certified, blacken the "N" block with a not too soft lead pencil.

	N	0	1
1	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

b. If the soldier does not qualify on the scorable unit, blacken the 0 block.

	N	0	1
1	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

c. If the soldier qualifies on the scorable unit, blacken the 1 block.

	N	0	1
1	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## APPENDIX A (continued)

SQT 2, MOSC 16J10

PERFORMANCE CERTIFICATION SCORESHEET  
(REFER TO INSTRUCTIONS TO SUPERVISORS)

SOLDIER'S NAME \_\_\_\_\_

SSN \_\_\_\_\_ UNIT OF ASSIGNMENT \_\_\_\_\_

PASS FAIL

1. At the radar power supply: Adjusts VOLTMETER to indicate between 110 and 120 volts.

2. Adjusts LINE FREQ meter to indicate between 384 and 416 CPS.

(NOTE: Sequence of these first two steps is NOT important.)

3. At the indicator: Turns POWER switch ON.

4. At the radar power supply: Looks at VOLTMETER and LINE FREQ meter and readjust if either is outside the limits specified in steps 1 and 2.

5. Blower motors should be heard operating in the POWER SUPPLY, RECEIVER-TRANSMITTER, and the ANTENNA BASE units. If one of these motors does not come on, and examinee proceeds, he fails this step.

6. At the indicator: Looks at or moves to and looks at the PPI-scope and A-scope. A spot of light should be visible on each scope.

7. At the receiver-transmitter: Positions TEST SELECTOR switch to MAG-FIL and looks at test meter. Meter should indicate between limits shown on stick-on template.

8. At the indicator: Turns ANTENNA ROTATE control clockwise and then counterclockwise. If antenna indicator fails to move, this constitutes an equipment malfunction.

(NOTE: At this point examinee must wait until delay timer runs down.)

9. At the indicator: Presses RADIATE pushbutton.

10. Performs steps in sequence.

11. Completes task in 10 minutes.

GO	NO GO	NOT OBSERVED
----	-------	-----------------

The soldier is scored GO if he passes all performance measures.

APPENDIX B

SQT NOTICE

1. PARAGRAPHS 1-3 will be the same on all notices.
2. PARAGRAPH 4 will be the written component.
  - a. Paragraph 4a is standard except for changes to the number of scorable units and MOS and skill level.
  - b. Paragraph 4b is specific to MOS and skill level.
3. PARAGRAPH 5 will be the hands-on portion if your test has a hands-on portion.
  - a. Paragraph 5a is standard.
  - b. Paragraph 5b is specific to the MOS and skill level. It will be transcribed from the instructions to examinee and the score sheets in the hands-on package.
4. PARAGRAPH 6 will be the performance certification component if your test has a PCC. If there is no hands-on, the PCC will be paragraph 5.
  - a. Paragraph a is a description of the purpose of the component and a brief verbal description of each task.
  - b. Paragraph b is an extract of the scoresheet from the performance certification package with a standard introduction which must be modified to match the skill level and MOS title.
5. HANDS-ON COMPONENT.
  - a. The second part of your SQT will be the hands-on component which will consist of \_\_\_\_\_ performance tests. Each test counts one scorable unit. You will be graded GO or NO GO by a scorer as you perform each of the \_\_\_\_\_ tests.
  - b. The tasks which will be included in the hands-on component are listed below. The task numbers and titles are taken from the Soldier's Manuals for Skill Levels \_\_\_\_\_ and \_\_\_\_\_, (Title of MOSC)

<u>TASK NUMBER</u>	<u>TASK</u>
54-1 000-00A-1003	PERFORM MASKING AND UNMASKING PROCEDURES

INSTRUCTIONS TO EXAMINEE:

AT THIS STATION YOU WILL BE TESTED ON YOUR ABILITY TO MASK PROPERLY WITH THE M17A1 MASK. WHEN I GIVE YOU THE ALERT "GAS," YOU WILL PUT ON THE MASK USING PROPER PROCEDURES. YOU HAVE 9 SECONDS TO COMPLETE THIS TASK. DO YOU UNDERSTAND WHAT YOU ARE REQUIRED TO DO? "GAS!"

PASS      FAIL

1. Puts on and seats protective mask by inserting chin first.
  2. Clears protective mask by placing one hand over the outlet cover and blowing.
  3. Checks seal by placing hands over air inlet valve assemblies and inhaling. Mask will collapse.
  4. Completes all steps within 9 seconds of the command "GAS."
  5. Perform all steps in sequence.
6. PERFORMANCE CERTIFICATION COMPONENT.
- a. The purpose of the performance certification component SQT is to allow your commander or supervisor to certify your proficiency with the M16A1 rifle and your physical fitness as measured by the Advanced Physical Fitness Test (APFT). You are required to take the APFT and qualify with the M16A1 rifle within the 12 months preceding the SQT quarter. Your performance will be certified on the following tasks as explained below.
  - b. The tasks which will be included in the performance certification are listed below. The task numbers and titles are taken from the Soldier's Manuals for Skill Levels and, (Title of MOSC).

<u>TASK NUMBER</u>	<u>TASK</u>	<u>SCORABLE UNITS</u>
SU-(1) 000-00A-2002	QUALIFY WITH M161 RIFLE (ENGAGE TARGETS WITH M16AI RIFLE) (SPECIFIC COURSE OF FIRE TO BE DETERMINE BY LOCAL CONDITIONS)	Up to 3 units a. Fail to Qualify: 0 units b. Marksman: 1 unit c. Sharpshooter: 2 units d. Expert: 3 units
SU-(2) 000-00A-2005	TAKE APFT (ADVANCED PHYSICAL FITNESS TEST) (MAINTAIN INDIVIDUAL PHYSICAL FITNESS)	Up to 3 units a. Fail to Qualify: 0 Units b. 300-399: 1 unit c. 400-499: 2 units d. 450-500: 3 units

## CHECKLIST FOR WRITTEN SU

## SQT NOTICE ENTRY

1. Verify and list the correct scorable unit number from the test.
2. Verify and list the correct task number from the Soldier's Manual.
3. Verify and list the correct task title from the Soldier's Manual.
4. Tell the soldier what he will be given.
  - a. format or problem (situation, question)
  - b. media (graphs, photos, tables, charts)
  - c. tools/aids (maps, protractors, JPA, manual extracts)
5. Tell the soldier what aspects or areas of the task he will be dealing with.
6. Tell the soldier what specific products he must select.
7. Tell the soldier how he will be required to select his answers (correct alternatives, sequencing, matching).
8. Tell the soldier how many questions he will be required to answer and how many he must answer correctly for a GO.
9. Show the soldier a sample item from the SU.

SAMPLE

SKILL QUALIFICATION TEST NOTICE  
FOR  
SQT\_\_\_\_\_, MOSC\_\_\_\_\_

1. THE SKILL QUALIFICATION TEST (SQT). The SQT is designed to test your ability to do those tasks in your Military Occupational Specialty (MOS) that are most important to your survival in combat, accomplishment of your job, and your unit's mission. The SQT does not test all these important tasks in any one year, and the sample may change each year.
2. SQT COMPONENTS. The SQT may have as many as three parts (components) or as few as one: Written Component, Hands-on Component, and Performance Certification Component. Each component is made up of Scorable Units. A Scorable Unit measures your ability to perform a task. The tasks included in the SQT are taken from the Soldier's Manual (SM) for the MOS being tested. Each SQT is made up of tasks taken from two skill levels. Most of the tasks come from the SM for the skill level in which you are working when you take the SQT. Some come from the SM for the next higher skill level. The components included in this SQT are described in paragraphs \_\_\_\_\_ below.
3. PREPARING TO TAKE THE SQT. In preparing to take your SQT, you should use your Soldier's Manual. For each task tested on your SQT, the Soldier's Manual tells you what you have to do to perform the task, the conditions under which the task is to be performed, and the standards which the performance has to meet. The task number and title of each task on the SQT is given below. Locate each task in the Soldier's Manual. Carefully read the description. Get the TM's, FM's, or other materials referred to in the Soldier's Manual and carefully study the sections which are important. Be sure you understand what is required to perform the task. If you have questions, ask your supervisor. When you clearly understand what the task is and what you are to do in performing the task, practice the task. Get your supervisor or someone else to score your performance. Continue the practice until you have mastered the task.
4. WRITTEN COMPONENT.

- a. Description. There will be \_\_\_\_\_ scorable units in the Written Component. Each scorable unit will have one to ten questions. You may be tested in several different ways in the Written Component. You may have to work through a problem and then answer a series of questions about what you have done. You may answer questions based on a situation or drawing. You may have to answer questions that show you know how to perform an action. All questions will be based on on-the-job performance.

SAMPLE

There will not be any trick questions. The best way to study and prepare for the SQT is to practice the tasks according to the Soldier's Manual.

b. Tasks. The tasks which will be included in the Written Component are listed below. The task numbers and titles are taken from the Soldier's Manuals for Skill Levels \_\_\_\_ and \_\_\_\_\_. (Title of MOSC.)

<u>TASK NUMBER</u>	<u>TASK</u>
SU-(1) 000-00A-0000	DETERMINE FIRST AID MEASURES TO BE APPLIED TO A CASUALTY

You will be given two different situations and will have to answer questions about each situation. The questions will lead you through a series of events which duplicates the normal treatment sequence for each type of casualty. You must choose the correct treatment or treatments from the possible answers provided. This unit will consist of four questions. You must answer all questions correctly to score a "GO" for the unit.

<u>TASK NUMBER</u>	<u>TASK</u>
SU-(2) 000-00A-0001	RECOGNIZE AND INTERPRET MARKERS OF AREAS OF NBC CONTAMINATION AND HAZARD

You will be given six drawings of markings with basic color information identified. You must match the correct shape and color with the words that complete the marker. This unit will consist of four questions. You must answer all questions correctly to score a "GO" for the unit.

<u>TASK NUMBER</u>	<u>TASK</u>
SU-(3) 000-00A-0002	INSPECT AND MAINTAIN THE PROTECTIVE MASK AND ITS CARRIER

You will be asked questions on inspection procedures, actions authorized for operator maintenance, cleaning, and repair procedures. You must select the correct action or procedure from the alternatives given. This unit will consist of five questions. You must answer four questions correctly to score a "GO" for the unit.

TASK NUMBERTASK

SU-(4) 000-00A-0003

PERFORM BATTLESIGHT ZEROING OF M16A1 RIFLE

You will be given pictures showing steps in zeroing the M16 rifle. You will have to answer two types of questions. In one type you have to identify the problem shown in the picture. In the other, you must answer the question by selecting the correct picture. This unit will consist of five questions. You must answer four questions correctly to score a "GO" for the unit.

TASK NUMBERTASK

SU-(5) 000-00A-004

DETERMINE YOUR LOCATION ON A MAP

You will be given a map protractor, grid scales, and selected parts of maps. Your position on the map or the coordinates and the azimuth will be given to you in the question. You must complete the necessary calculations or conversions asked for and select the correct answer. This unit will consist of four questions. You must answer four questions correctly to score a "GO" for the unit.

TASK NUMBERTASK

SU-(6) 000-00A-0009

PERFORM EQUIPMENT SERVICEABILITY CRITERIA (ESC) INSPECTION.

You will be given a portion of an Equipment Serviceability Criteria (ESC) and asked questions about the performance of an ESC inspection. You will be given also a symptom in the question and be required to identify the correct status to place on the part of the equipment that has the symptom. You must select the correct answer from those given. This unit will consist of three questions. You must answer three questions correctly to score a "GO" for the unit.

TASK NUMBERTASK

SU-(7) 00-00A-011

PREPARE TAMMS FORMS USED BY VEHICLE OPERATOR

You will be given a partially completed DA Form 2400 (Equipment Utilization Record) and a blank DA Form 2404 (Equipment Inspection/Maintenance Worksheet). In the situation, you will be assigned as a driver of a vehicle and asked to identify where entries are made on the DA Form 2400 and 2404 or asked questions about the use of the forms. You must choose the correct answers from sets of possible answers. This unit will consist of three questions. You must answer three questions correctly to score a "GO" for the unit.

GO NO GO

## UNIT 1

<u>TASK NUMBER</u>	<u>TASK</u>
54-2 00-00A-1002	INSTALL AND OPERATE FIELD TELEPHONE TA-312/PT

INSTRUCTIONS TO THE EXAMINEE: YOU WILL INSTALL AND OPERATE A TA-312/PT, FIELD TELEPHONE. THE FIELD WIRE YOU SEE HERE IS CONNECTED TO ANOTHER FIELD TELEPHONE WHICH WILL ACT AS A SWITCHBOARD FOR YOU. ALL EQUIPMENT NEEDED TO INSTALL AND OPERATE THE TA-312/PT IS PRESENT. I CANNOT HELP YOU DURING THIS TEST. YOU HAVE FIVE MINUTES TO COMPLETE THIS TASK. DO YOU UNDERSTAND WHAT YOU ARE REQUIRED TO DO? YOU MAY BEGIN.

	<u>PASS</u>	<u>FAIL</u>
1. Uses pliers TL-13A to strip insulation from each conductor of the wire line.	—	—
2. Presses spring-loaded line binding posts and inserts one conductor into each post.	—	—
3. Adjusts buzzer volume control knob to LOUD position.	—	—
4. Places INT-EXT switch to INT.	—	—
5. Places CB/LB/CBS switch to LB.	—	—
6. Inserts BA-30 batteries into the battery compartment, one up and one down.	—	—
"NOW RING THE OPERATOR ON THE OTHER END AND ASK HIM TO RING YOU BACK."		
7. Turns handcrank rapidly at least two turns.	—	—
8. Removes handset from retaining cradle and waits for operator to answer.	—	—
9. Presses PRESS-TO-TALK switch and asks operator on other end to ring back.	—	—
10. When buzzer rings, terminates call and places handset in retaining cradle.	—	—
11. Performs first six steps before step seven.	—	—
12. Completes task in five minutes.	—	—

GO NO GO

## UNIT 2